Use of DNA Encoding Plastid Pyruvate Dehydrogenase and Branched Chain Oxoacid Dehydrogenase Components to Enhance Polyhydroxyalkanoate Biosynthesis in Plants

This application claims the benefit of priority of the following Provisional patent applications: Serial Number 60/051,291, filed June 30, 1997; Serial Number 60/055,255, filed August 1, 1997; Serial Number 60/076,544, filed March 2, 1998; and Serial Number 60/076,554, filed March 2, 1998.

Background of the Invention

Field of the Invention

The present invention relates to genetically engineered plants. More particularly, the present invention relates to the optimization of substrate pools to facilitate the biosynthetic production of commercially useful polyhydroxyalkanoates (PHAs) in plants.

The present invention especially relates to the production of copolyesters of 3-hydroxybutyrate (3HB) and 3-hydroxyvalerate (3HV), designated P(3HB-co-3HV) copolymer, and derivatives thereof.

Description of Related Art

Polyhydroxyalkanoates

Polyhydroxyalkanoates are polyesters that accumulate in a wide variety of bacteria. These polymers have properties ranging from stiff and brittle plastics to

rubber-like materials, and are biodegradable. Due to these properties, PHAs are an attractive source of non-polluting plastics and elastomers.

Currently, there are approximately a dozen biodegradable plastics in commercial use that possess properties suitable for producing a number of specialty and commodity products (Lindsay, 1992). One such biodegradable plastic in the polyhydroxyalkanoate (PHA) family that is commercially important is Biopol™, a random copolymer of 3-hydroxybutyrate (3HB) and 3-hydroxyvalerate (3HV). This bioplastic is used to produce biodegradable molded material (e.g., bottles), films, coatings, and in drug release applications. Biopol™ is produced via a fermentation process employing the bacterium *Alcaligenes eutrophus* (Byrom, 1987). The current market price is $6-7/lb, and the annual production is 1,000 tons. By best estimates, this price is likely to be reduced only about 2-fold via fermentation (Poirier et al., 1995). Competitive synthetic plastics such as polypropylene and polyethylene cost about 35-45¢/lb (Layman, 1994). The annual global demand for polyethylene alone is about 37 million metric tons (Poirier et al., 1995). It is therefore likely that the cost of producing P(3HB-co-3HV) by microbial fermentation will restrict its use to low-volume specialty applications.

Nakamura et al. (1992) reported using threonine (20g/L) as the sole carbon source for the production of P(3HB-co-3HV) copolymer in *A. eutrophus*. These workers initially suggested that the copolymer might form via the degradation of threonine by threonine deaminase, with

conversion of the resultant α-ketobutyrate (= 2-oxobutyrate) to propionyl-CoA. However, they ultimately concluded that threonine was utilized directly, without breaking carbon-carbon bonds, to form valeryl-CoA as the 3HV precursor. The nature of this chemical conversion was not described, but since the breaking of carbon-carbon bonds was not postulated to occur, the pathway could not involve threonine deaminase in conjunction with an α-ketoacid decarboxylating step to form propionate or propionyl-CoA. In the experiments of Nakamura et al., the PHA polymer content was very low (< 6% of dry cell weight). This result, in conjunction with the expense of feeding bacteria threonine, makes their approach impractical for the commercial production of P(3HB-co3HV) copolymer.

Yoon et al. (1995) have shown that growth of *Alcaligenes* sp. SH-69 on a medium supplemented with threonine, isoleucine, or valine resulted in significant increases in the 3HV fraction of the P(3HB-co-3HV) copolymer. In addition to these amino acids, glucose (3% wt/vol) was also added to the growth media. In contrast to the results obtained by Nakamura et al. (1992), growth of *A. eutrophus* under the conditions described by Yoon et al. (1995) did not result in the production of P(3HB-co-3HV) copolymer when the medium was supplemented with threonine as the sole carbon source. From their results, Yoon et al. (1995) implied that the synthetic pathway for the 3HV component in P(3HB-co-3HV) copolymer is likely the same as that described in WO 91/18995 and Steinbüchel and Pieper (1992). This postulated synthetic pathway involves the degradation of isoleucine to

propionyl-CoA (Figure 3).

The PHB Biosynthetic Pathway

Polyhydroxybutyrate (PHB) was first discovered in 1926 as a constituent of the bacterium *Bacillus megaterium* (Lemoigne, 1926). Since then, PHAs such as PHB have been found in more than 90 different genera of gram-negative and gram-positive bacteria (Steinbüchel, 1991). These microorganisms produce PHAs using R-β-hydroxyacyl-CoAs as the direct metabolic substrate for a PHA synthase, and produce polymers of R-(3)-hydroxyalkanoates having chain lengths ranging from C3-C14 (Steinbüchel and Valentin, 1995).

To date, the best understood biochemical pathway for PHB production is that found in the bacterium *Alcaligenes eutrophus* (Dawes and Senior, 1973; Slater et al., 1988; Schubert et al., 1988; Peoples and Sinskey, 1989a and 1989b). This pathway, which is also utilized by other microorganisms, is summarized in Figure 1. In this organism, an operon encoding three gene products, i.e., PHB synthase, β-ketothiolase, and acetoacetyl-CoA reductase, encoded by the *phbC*, *phbA*, and *phbB* genes, respectively, are required to produce the PHA homopolymer R-polyhydroxy-butyrate (PHB).

As further shown in Figure 1, acetyl-CoA is the starting substrate employed in the biosynthetic pathway. This metabolite is naturally available for PHB production in the cytoplasm and plastids of plants.

Poirier et al. (1992) demonstrated that a multi-enzyme pathway can be introduced into plants to produce polyhydroxybutyrate (PHB). In that work, the genes

encoding the *Alcaligenes eutrophus* acetoacetyl-CoA reductase (*phbB*) and PHB synthase (*phbC*) genes were introduced into *Arabidopsis thaliana*, where the enzymes were expressed cytoplasmically. A 3-ketothiolase is already expressed in the cytoplasm of *Arabidopsis*. Although PHB was produced in the plants which expressed the three enzymes, the yield was low and the plants were stunted and had reduced seed production.

Nawrath et al. (1994) provided a solution to these problems. There, the genes for the three bacterial PHB enzymes (*phbC*, *phbA*, and *phbB*) were modified to comprise a pea chloroplast targeting peptide (="transit peptide"), which targeted the enzymes to the chloroplast. *Arabidopsis* plants which produced these three enzymes in the chloroplast accumulated large amounts of PHB. There was also no apparent affect of these transgenes, or of the PHB accumulation, on the growth and development of the transgenic plants.

The P(3HB-co-3HV) Copolymer Biosynthetic Pathway

As noted above, P(3HB-co-3HV) random copolymer, commercially known as Biopol™, is produced by fermentation employing *A. eutrophus*. A proposed biosynthetic pathway for P(3HB-co-3HV) copolymer production is shown in Figure 2. Production of this polymer in plants has been reported (oral presentation by Mitsky et al., 1997).

Since the production of PHB in chloroplasts apparently does not affect plant growth and development as does production of PHB in the cytoplasm (Nawrath et al., 1992), the chloroplast is the preferred site of

P(3HB-co-3HV) biosynthesis. The successful production of P(3HB-co-3HV) copolymer in plants thus requires the presence of three PHA biosynthetic enzymes as well as the substrates required for the copolymer biosynthesis (Figure 2), preferably in the plastids. For the 3HB component of the polymer, the substrate naturally exists in chloroplasts in sufficient concentration in the form of acetyl-CoA (Nawrath et al., 1994). However, this is not true for the 3HV component of the polymer, where the starting substrate is propionyl-CoA. Figure 3 is an overview of enzyme pathways which are related to the provision of these substrates. The engineering of plants to generate sufficient chloroplast pools of propionyl-CoA, along with the proper PHA biosynthetic enzymes (i.e., a β-ketothiolase, a β-ketoacyl-CoA reductase, and a PHA synthase), makes it possible to produce copolyesters of poly(3HB-co-3HV) in these organisms.

Methods for optimization of PHB and P(3HB-co-3HV) production in various crop plants are disclosed in Gruys et al. (1998). A major focus in that invention is the optimization of the substrate pools for P(3HB-co-3HV), in order to provide 2-ketobutyrate and propionyl-CoA to the site of copolymer synthesis. Gruys et al. (1998) also discusses exploring the potential use of a pyruvate dehydrogenase complex and a branched chain oxoacid dehydrogenase complex to convert 2-oxobutyrate to propionyl-CoA.

Gruys et al. (1998) also provides methods for the optimization of β-ketothiolase, β-ketoacyl-CoA reductase, and PHA synthase activities in plants and bacteria. It was determined therein that the *A. eutrophus* β-

ketothiolase PhbB was metabolically blocked from producing P(3HB-co-3HV) due to its inability to utilize propionyl-CoA with acetyl-CoA to produce 3-ketovaleryl-CoA (see Figure 2). However, Gruys et al. (1998) demonstrated that another *A. eutrophus* β-ketothiolase, designated BktB, is able to produce 3-ketovaleryl-CoA from propionyl-CoA and acetyl-CoA. Therefore, BktB is a preferred β-ketothiolase for the production of P(3HB-co-3HV). Gruys et al. also demonstrated that other β-ketothiolases are able to produce 3-keto-valeryl-CoA from propionyl-CoA and acetyl-CoA. These are: another *A. eutrophus* β-ketothiolase, designated pAE65, and two β-ketothiolases from *Zoogloea ramigera*, designated "A" and "B".

Gruys et al. (1998) noted that the sources of the three copolymer biosynthetic enzymes may encompass a wide range of organisms, including, for example, *Alcaligenes eutrophus*, *Alcaligenes faecalis*, *Aphanothece* sp., *Azotobacter vinelandii*, *Bacillus cereus*, *Bacillus megaterium*, *Beijerinkia indica*, *Derxia gummosa*, *Methylobacterium* sp., *Microcoleus* sp., *Nocardia corallina*, *Pseudomonas cepacia*, *Pseudomonas extorquens*, *Pseudomonas oleovorans*, *Rhodobacter sphaeroides*, *Rhodobacter capsulatus*, *Rhodospirillum rubrum* (Brandl et al., 1990; Doi, 1990), and *Thiocapsa pfennigii*.

Pyruvate Dehydrogenase Complex

The pyruvate dehydrogenase complex (PDC) is a large multi-enzyme structure composed of three primary component enzymes, pyruvate dehydrogenase (PDH) (E1, EC 1.2.41); dihydrolipoamide acetyltransferase (E2, EC

2.3.1.12); and dihydrolipoamide dehydrogenase (E3, EC 1.8.1.4) (Reed, 1974). In the well-characterized mammalian complex, 60 subunits of E2 comprise the central core, and the E1 and E3 components decorate the outer surface of this core (Patel et al., 1990). E1 is a heterotetramer composed of two α and two β subunits. The E3 component, a homodimer, associates with the complex via an E3 binding protein (Gopalakrishnan, 1989).

The PDC catalyzes the irreversible oxidative decarboxylation of pyruvate according to the equation:

$$\text{Pyruvate} + \text{CoA} + \text{NAD}^+ \rightarrow \text{Acetyl-CoA} + \text{CO}_2 + \text{NADH} + \text{H}^+$$

In mitochondria, this reaction represents the irreversible commitment of carbon to the citric acid cycle, and therefore is a logical point for regulation. Previous experiments have shown that plant mitochondrial PDC activity is, in fact, regulated by product inhibition, metabolites, and reversible phosphorylation (Randall et al., 1977; Randall et al., 1989; Randall et al., 1996; Budde et al, 1991) as is the mammalian complex (Patel et al., 1990).

In prokaryotes, PDC is localized in the cytoplasm, while in eukaryotes it is within the mitochondrial matrix. Plants, however, are unique in that a second form of the complex exists in the plastids (Reid et al., 1975; Reid et al., 1977; Thompson et al, 1977b). Based upon enzymology (Thompson et al., 1977a; Williams et al., 1979; Camp et al., 1988) and immunochemical analyses (Taylor et al., 1992; Camp et al, 1985) it is clear that plastid PDC is distinct from its mitochondrial

counterpart. In plants, *de novo* fatty acid biosynthesis occurs exclusively in the plastids (Miernyk et al., 1983; Kang et al., 1994; Zilket et al., 1969; Drennan et al., 1969; Ohlrogge et al., 1979). The plastid form of PDC can provide the fatty acid precursor, acetyl-CoA (Miernyk et al., 1983; Kang et al., 1994; Grof et al., 1995). The plastid PDC can also catalyze the oxidative decarboxylation of 2-oxobutyrate to produce propionyl Co-A (Camp et al., 1988; Camp and Randall, 1985).

The cDNAs that encode the E1α and E1β subunits of plant mitochondrial PDH have been cloned (Grof et al., 1995; Leuthy et al., 1995; Leuthy et al, 1994). Recently, Reith and Munholland (1995) reported the sequence of the entire plastid genome of the red alga *P. purpurea*. Encoded in this genome are open reading frames homologous to PDH α and β subunits.

The cDNAs that encode the E2 component of the plant mitochondrial PDC have been similarly cloned (Guan et al., 1995). The sequence of the entire plastid genome of the cyanobacterium *Synechocystis* sp. has also recently been reported (Kaneko et al., 1996).

Branched Chain 2-Oxoacid Dehydrogenase Complex

The branched chain 2-oxoacid dehydrogenase complex (BCOADC) is a highly ordered macromolecular structure composed of three primary component enzymes, a branched chain dehydrogenase or decarboxylase (BCDH or E1; EC 1.2.4.4); dihydrolipoamide transacylase (LTA or E2; no EC number); and dihydrolipoamide dehydrogenase (LipDH or E3;

EC 1.8.1.4) (Yeaman, 1989). The mammalian complex is assembled with 24 subunits of E2 as the central cubic core with 4:3:2 symmetry; the E1 and E3 components decorate the outer surface of the E2 core (Yeaman, 1989; Wynn et al., 1996). E1 is a heterotetramer composed of two identical α and two identical β subunits (Pettit et al., 1978). E3 associates loosely with the E2-E1 structure, and is a homodimer (Wynn et al., 1996; Pettit et al., 1978). The mammalian mitochondrial complex is also regulated by a specific E1-kinase and a phospho-E1 phosphatase, which modulate activity by reversible phosphorylation (inactivation) and dephosphorylation (reactivation). Additional regulation is achieved by product inhibition and modulation of gene expression (Yeaman, 1989; Wynn et al., 1996).

BCOADC catalyzes the irreversible oxidative decarboxylation of the branched-chain 2-oxoacids derived from valine, leucine and isoleucine, as well as 2-oxobutyrate and 4-methyl-2-oxobutyrate, with comparable rates and similar Km values (Yeaman 1989; Wynn et al., 1996; Paxton et al., 1986; Gerbling et al., 1988). The reactions are:

2-oxo-3-methylvalerate + CoA + NAD^+ → 2-methylbutyryl-CoA + CO_2 + NADH + H^+

2-oxo-isovalerate + CoA + NAD^+ → isobutyryl-CoA + CO_2 + NADH + H^+

2-oxo-isocaproiate + CoA + NAD^+ → isovalyryl-CoA + CO_2 + NADH + H^+

2-oxobutyrate + CoA + NAD^+ → propionyl-CoA + CO_2 + NADH + H^+

In mammals, BCOADC is found in the mitochondria and is involved in the catabolism of the branched-chain amino acids. The only reports describing BCOADC activity in

plants have localized BCOADC to peroxisomes (Gerbling et al., 1988; Gerbling et al., 1989). The proposed function of a peroxisomal BCOADC is to catabolize the branched-chain amino acids during germination and growth, yielding an acyl-CoA product that would be further metabolized by the beta-oxidation pathway localized in the peroxisome (Gerbling et al., 1988; Gerbling et al., 1989).

To provide substrate pools to permit biosynthesis of P(3HB-co-3HV) copolymer in the plastid, there is a need for methods to engineer plants to produce plastid enzymes which convert 2-oxobutyrate to propionyl-CoA.

Summary of the Invention

Accordingly, the present invention provides nucleotide sequences that encode the E1α and E1β subunits, and the E2 component of the plastid pyruvate dehydrogenase complex, as well as the E1α and E1β subunits, and the E2 component of the branched chain oxoacid dehydrogenase complex, of *Arabidopsis thaliana*. Methods of utilizing these nucleotide sequences to provide enzymatic activity to convert 2-oxo-butyrate to propionyl-CoA, and to produce P(3HB-co-3HV) copolymer in plants, are also provided.

Accordingly, in a first aspect, the present invention provides an isolated DNA molecule, comprising a nucleotide sequence selected from: (a) the nucleotide sequence shown in SEQ ID NO:1, or the complement thereof; (b) a nucleotide sequence that hybridizes to the nucleotide sequence of (a) under a wash stringency equivalent to 0.5X SSC to 2X SSC, 0.1% SDS, at 55-65°C, and which encodes a polypeptide having enzymatic activity

similar to that of *Arabidopsis thaliana* plastid pyruvate dehydrogenase complex E1α subunit; (c) a nucleotide sequence encoding the same genetic information as the nucleotide sequence of (a), but which is degenerate in accordance with the degeneracy of the genetic code; and (d) a nucleotide sequence encoding the same genetic information as the nucleotide sequence of (b), but which is degenerate in accordance with the degeneracy of the genetic code. Recombinant vectors comprising such isolated DNA molecules, host cells transformed with these vectors, and an isolated polypeptide having the amino acid sequence of SEQ ID NO.:2 are also provided.

In another aspect, the present invention provides an isolated DNA molecule, comprising a nucleotide sequence selected from: (a) the nucleotide sequence shown in SEQ ID NO:3, or the complement thereof; (b) a nucleotide sequence that hybridizes to the nucleotide sequence of (a) under a wash stringency equivalent to 0.5X SSC to 2X SSC, 0.1% SDS, at 55-65°C, and which encodes a polypeptide having enzymatic activity similar to that of *Arabidopsis thaliana* plastid pyruvate dehydrogenase complex E1β subunit; (c) a nucleotide sequence encoding the same genetic information as the nucleotide sequence of (a), but which is degenerate in accordance with the degeneracy of the genetic code; and (d) a nucleotide sequence encoding the same genetic information as the nucleotide sequence of (b), but which is degenerate in accordance with the degeneracy of the genetic code. Recombinant vectors comprising such isolated DNA molecules, host cells transformed with these vectors, and an isolated polypeptide having the amino acid sequence of

SEQ ID NO.:4 are also provided.

In another aspect, the present invention provides an isolated DNA molecule, comprising a nucleotide sequence selected from: (a) the nucleotide sequence shown in SEQ ID NO:5, or the complement thereof; (b) a nucleotide sequence that hybridizes to the nucleotide sequence of (a) under a wash stringency equivalent to 0.5X SSC to 2X SSC, 0.1% SDS, at 55-65°C, and which encodes a polypeptide having enzymatic activity similar to that of *Arabidopsis thaliana* plastid pyruvate dehydrogenase complex E2 component; (c) a nucleotide sequence encoding the same genetic information as the nucleotide sequence of (a), but which is degenerate in accordance with the degeneracy of the genetic code; and (d) a nucleotide sequence encoding the same genetic information as the nucleotide sequence of (b), but which is degenerate in accordance with the degeneracy of the genetic code. Recombinant vectors comprising such isolated DNA molecules, host cells transformed with these vectors, and an isolated polypeptide having the amino acid sequence of SEQ ID NO.:6 are also provided.

In a further aspect, the present invention provides an isolated DNA molecule, comprising a nucleotide sequence selected from: (a) the nucleotide sequence shown in SEQ ID NO:11, or the complement thereof; (b) a nucleotide sequence that hybridizes to the nucleotide sequence of (a) under a wash stringency equivalent to 0.5X SSC to 2X SSC, 0.1% SDS, at 55-65°C, and which encodes a polypeptide having enzymatic activity similar to that of *Arabidopsis thaliana* branched chain 2-oxoacid dehydrogenase complex E1α subunit; (c) a nucleotide

sequence encoding the same genetic information as the nucleotide sequence of (a), but which is degenerate in accordance with the degeneracy of the genetic code; and (d) a nucleotide sequence encoding the same genetic information as the nucleotide sequence of (b), but which is degenerate in accordance with the degeneracy of the genetic code. Recombinant vectors comprising such isolated DNA molecules, host cells transformed with these vectors, and an isolated polypeptide having the amino acid sequence of SEQ ID NO.:12 are also provided.

In another aspect, the present invention provides an isolated DNA molecule, comprising a nucleotide sequence selected from: (a) the nucleotide sequence shown in SEQ ID NO:13, or the complement thereof; (b) a nucleotide sequence that hybridizes to the nucleotide sequence of (a) under a wash stringency equivalent to 0.5X SSC to 2X SSC, 0.1% SDS, at 55-65°C, and which encodes a polypeptide having enzymatic activity similar to that of *Arabidopsis thaliana* branched chain 2-oxoacid dehydrogenase complex E1β subunit; (c) a nucleotide sequence encoding the same genetic information as the nucleotide sequence of (a), but which is degenerate in accordance with the degeneracy of the genetic code; and (d) a nucleotide sequence encoding the same genetic information as the nucleotide sequence of (b), but which is degenerate in accordance with the degeneracy of the genetic code. Recombinant vectors comprising such isolated DNA molecules, host cells transformed with these vectors, and an isolated polypeptide having the amino acid sequence of SEQ ID NO.:14 are also provided.

In another aspect, the present invention provides

the foregoing isolated DNA molecules encoding a polypeptide having enzymatic activity similar to that of *Arabidopsis thaliana* branched chain 2-oxoacid dehydrogenase complex E1β subunit, but in which the naturally occurring branched chain oxoacid dehydrogenase complex E2 component binding region thereof is replaced with the E2 component binding region of a plastid pyruvate dehydrogenase complex E1β subunit. The plastid pyruvate dehydrogenase complex E1β subunit can have the sequence shown in SEQ ID NO.:3. Recombinant vectors comprising such isolated DNA molecules, host cells transformed with these vectors, and the isolated polypeptide are also provided.

In yet another aspect, the present invention provides an isolated DNA molecule, comprising a nucleotide sequence selected from: (a) the nucleotide sequence shown in SEQ ID NO:15, or the complement thereof; (b) a nucleotide sequence that hybridizes to the nucleotide sequence of (a) under a wash stringency equivalent to 0.5X SSC to 2X SSC, 0.1% SDS, at 55-65°C, and which encodes a polypeptide having enzymatic activity similar to that of *Arabidopsis thaliana* branched chain 2-oxoacid dehydrogenase complex E2 component; (c) a nucleotide sequence encoding the same genetic information as the nucleotide sequence of (a), but which is degenerate in accordance with the degeneracy of the genetic code; and (d) a nucleotide sequence encoding the same genetic information as the nucleotide sequence of (b), but which is degenerate in accordance with the degeneracy of the genetic code. Recombinant vectors comprising such isolated DNA molecules, host cells

transformed with these vectors, and an isolated polypeptide having the amino acid sequence of SEQ ID NO.:16 are also provided.

In another aspect, the present invention provides a plant, a plastid of which comprises the following polypeptides: an enzyme that enhances the biosynthesis of 2-oxobutyrate; a branched chain oxoacid dehydrogenase complex E1α subunit; a branched chain oxoacid dehydrogenase complex E1β subunit; and a branched chain oxoacid dehydrogenase complex E2 component. The branched chain oxoacid dehydrogenase complex E1α subunit can have the sequence shown in SEQ ID NO.:12, the branched chain oxoacid dehydrogenase complex E1β subunit can have the sequence shown in SEQ ID NO.:14, or the branched chain oxoacid dehydrogenase complex E2 component can have the sequence shown in SEQ ID NO.:16. In such plant, the plastid can further comprise the following polypeptides: a β-keto-thiolase; a β-ketoacyl-CoA reductase; and a polyhydroxy-alkanoate synthase. The genome of such plant can comprise introduced DNAs encoding these polypeptides, wherein each of the introduced DNAs is operatively linked to a targeting peptide coding region capable of directing transport of the polypeptide encoded thereby into a plastid. A method of producing P(3HB-co-3HV) copolymer comprises growing such plant, and recovering P(3HB-co-3HV) copolymer produced thereby.

In another aspect, the present invention comprises a plant, a plastid of which comprises the following polypeptides: an enzyme that enhances the biosynthesis of 2-oxobutyrate; a branched chain oxoacid dehydrogenase complex E1α subunit; a branched chain oxoacid

dehydrogenase complex E1β subunit; a branched chain oxoacid dehydrogenase complex E2 component; and a dihydrolipoamide dehydrogenase E3 component, which can be mitochondrially-derived. In such plant, the branched chain oxoacid dehydrogenase complex E1α subunit can have the sequence shown in SEQ ID NO.:12, the branched chain oxoacid dehydrogenase complex E1β subunit can have the sequence shown in SEQ ID NO.:14, or the branched chain oxoacid dehydrogenase complex E2 component can have the sequence shown in SEQ ID NO.:16. In such plant, the plastid can further comprise the following polypeptides: a β-keto-thiolase; a β-ketoacyl-CoA reductase; and a polyhydroxy-alkanoate synthase. The genome of such plant can comprise introduced DNAs encoding these polypeptides, wherein each of the introduced DNAs is operatively linked to a targeting peptide coding region capable of directing transport of the polypeptide encoded thereby into a plastid. A method of producing P(3HB-co-3HV) copolymer comprises growing such plant, and recovering P(3HB-co-3HV) copolymer produced thereby.

In yet another aspect, the present invention provides a plant, a plastid of which comprises the following polypeptides: an enzyme that enhances the biosynthesis of 2-oxobutyrate; a branched chain oxoacid dehydrogenase complex E1α subunit; and a branched chain oxoacid dehydrogenase complex E1β subunit, the naturally occurring E2 binding region of which is replaced with the E2 binding region of a plastid pyruvate dehydrogenase complex E1β subunit. In such plant, the branched chain oxoacid dehydrogenase complex E1α subunit can have the

sequence shown in SEQ ID NO.:12. Furthermore, in such plant, the plastid can further comprise the following polypeptides:

a β-ketothiolase; a β-ketoacyl-CoA reductase; and a poly-hydroxyalkanoate synthase. In such plant, the genome can comprise introduced DNAs encoding these polypeptides, wherein each of the introduced DNAs is operatively linked to a targeting peptide coding region capable of directing transport of the polypeptide encoded thereby into a plastid.

A method of producing P(3HB-co-3HV) copolymer comprises growing such plant, and recovering P(3HB-co-3HV) copolymer produced thereby.

Further scope of the applicability of the present invention will become apparent from the detailed description and drawings provided below. However, it should be understood that the following detailed description and examples, while indicating preferred embodiments of the invention, are given by way of illustration only since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

Brief Description of the Drawings

The above and other objects, features, and advantages of the present invention will be better understood from the following detailed description taken in conjunction with the accompanying drawings, all of which are given by way of illustration only, and are not limitative of the present invention, in which:

Figure 1 shows the biochemical steps involved in the production of PHB from acetyl-CoA catalyzed by the *A. eutrophus* PHB biosynthetic enzymes.

Figure 2 shows the biochemical steps involved in the production of P(3HB-co-3HV) copolymer from acetyl-CoA and propionyl-CoA catalyzed by PHA biosynthetic enzymes of *A. eutrophus*.

Figure 3 summarizes the pathways discussed herein that are involved in the production of P(3HB-co-3HV) copolymer, including enzymes that can be used to enhance 2-oxobutyrate biosynthesis.

Figure 4 shows Southern analyses of genomic DNA isolated from mature *A. thaliana* leaves. Each lane was loaded with 10 μg of DNA digested with *Bam*HI, *Hind* III, *Sal* I, *Eco* RI or *Xba* I as indicated. Fig. 5A and 5B, genomic Southern blots hybridized with random primed probes generated from gel-excised E1α and E1β cDNAs respectively. (α^{32}P)-dCTP was incorporated using an oligolabelling kit (Pharmacia, Uppsala, Sweden). The positions of λ DNA markers digested with *Hind* III are indicated to the left of the figure.

Figure 5 shows Northern blot analyses of *A. thaliana* RNA. Total RNA was isolated from young leaves of *A. thaliana* plants. 10 μg of total RNA was run on formaldehyde gels then transferred to nylon membranes. Probes were prepared as described in the legend for Figure 5. RNA markers were used to determine the sizes of the hybridizing bands.

Figures 6A and 6B show dendrogram analyses of the deduced amino acid sequence of PDH E1α and E1β subunits, respectively. Abbreviations and accession numbers to the

sequences are: *P. p., Porphyra purpurea odp* (U38804); *S. sp., Synechocystis sp.* (D90915); *A. t., Arabidopsis thaliana* (U21214, U09137); *P. s., Pisum sativum* (U51918, U56697); *H. s., Homo sapiens* (L13318, D90086); *R. r., Rattus rattus* (Z12158, P49432); *S. c., Saccharomyces cerevisiae* (P16387, M98476); *A. s., Ascaris suum* (M76554, M38017); *M. gen., Mycoplasma genetalium* (U39706); *M. c., Mycoplasma capricolum* (U62057); *B. su., Bacillus subtilis* (M57435); and *B. s., Bacillus stearothermophilus* (X53560). Dendrogram analyses was accomplished with GeneWorks CLUSTAL V method (IntelliGenetics, Mountain View, CA). CLUSTAL V parameters were as follows: cost to open gap = 5, cost to lengthen gap = 25, gap penalty = 3, number of top diagonals = 5, window size = 5, PAM matrix = PAM250, K-tuple = 1, and consensus cutoff = 50%.

Figures 7A-7E shows schematics (Constructs 1-5) for engineering the BCOADC subunits to be targeted to the plastid and to form a hybrid complex, as described in Examples 6 and 7.

IN 437

Detailed Description of the Invention

The following detailed description is provided to aid those skilled in the art in practicing the present invention. Even so, this detailed description should not be construed to unduly limit the present invention as modifications and variations in the embodiments discussed herein can be made by those of ordinary skill in the art without departing from the spirit or scope of the present inventive discovery.

The contents of each of the references cited herein, including those of the references cited within these primary references, are herein incorporated by reference in their entirety.

The production of P(3HB-co-3HV) in plants requires the substrates propionyl-CoA and acetyl-CoA, and three enzymes which convert these substrates to P(3HB-co-3HV): a β-ketothiolase, a β-ketoacyl-CoA reductase, and a PHA synthase. β-ketothiolase is normally present in the plant cytoplasm, but not in the plastids. Acetyl-CoA is normally present in the cytoplasm and the plastids. All of the other required components must be introduced into the plant, preferably into the plastids.

Gruys et al. (1998) discusses several ways in which 2-oxobutyrate can be provided in the plant. One way is through the manipulation of various wild-type and/or deregulated enzymes involved in the biosynthesis of aspartate family amino acids in order to increase threonine levels, thereby creating a larger substrate pool for threonine deaminase to convert to 2-oxobutyrate (Figure 3), and wild-type or deregulated forms of enzymes, such as threonine deaminase, involved in the conversion of threonine to P(3HB-co-3HV) copolymer endproduct. Enzymes which can be manipulated to enhance the threonine pool include aspartate kinase, homoserine dehydrogenase, and threonine synthase. The threonine substrate pool can be enhanced by overexpression of these enzymes, or by the use of deregulated forms of these enzymes, such as lysine-deregulated aspartate kinase.

Threonine deaminase, which converts threonine to 2-oxobutyrate, is another enzyme which can be utilized in

the production of 2-oxobutyrate. Deregulated mutants and natural deregulated forms of threonine deaminase can be used to increase 2-oxobutyrate pools at the site of copolymer biosynthesis.

Gruys et al. (1998), at Example 6, also discuss several ways in which the PDC and/or the BCOADC, or their substrate pools, can be manipulated to provide effective conversion of 2-oxobutyrate to propionyl-CoA. The native plastid PDC is able to perform this conversion at a low level. However, this complex can provide levels of propionyl-CoA sufficient for P(3HB-co-3HV) if the levels of 2-oxobutyrate are sufficient, or if portions of the BCOADC are employed to form a hybrid complex. The plastid PDC might also be genetically manipulated to be more effective in providing propionyl-CoA (Gruys et al., 1998).

The present invention provides nucleotide sequences that encode the E1α and E1β subunits, and the E2 component, of the plastid pyruvate dehydrogenase complex, and the E1α and E1β subunits, and the E2 component, of the branched chain oxoacid dehydrogenase complex of *Arabidopsis thaliana*. These nucleotide sequences and the enzymatic polypeptides encoded thereby can be introduced into plants in various combinations with coding sequences for the foregoing enzymes in order to enhance the conversion of threonine to 2-oxobutyrate, propionate, propionyl-CoA, β-ketovaleryl-CoA, and β-hydroxyvaleryl-CoA. Introduction into such plants of nucleic acid sequences encoding an appropriate β-keto-thiolase, a β-ketoacyl-CoA reductase, and a PHA

synthase will permit such transgenic plants to utilize the increased β-hydroxyvaleryl-CoA substrate in the production of
P(3HB-co-3HV) copolymer.

Definitions

The following definitions are provided to aid those skilled in the art in understanding the detailed description of the present invention.

"β-ketoacyl-CoA reductase" refers to a β-ketoacyl-CoA reducing enzyme that can convert a β-ketoacyl-CoA substrate to its corresponding β-hydroxyacyl-CoA product using, for example, NADH or NADPH as the reducing cosubstrate. An example is the PhbB acetoacetyl-CoA reductase of *A. eutrophus*.

"β-ketothiolase" refers to an enzyme that catalyzes the thiolytic cleavage of a β-ketoacyl-CoA, requiring free CoA, to form two acyl-CoA molecules. However, the term β-ketothiolase as used herein also refers to enzymes that catalyze the condensation of two acyl-CoA molecules to form β-ketoacyl-CoA and free CoA, i.e., the reverse of the thiolytic cleavage reaction.

"CoA" refers to coenzyme A.

"C-terminal" refers to the region of a peptide, polypeptide, or protein chain from the middle thereof to the end that carries the amino acid having a free α carboxyl group.

"Deregulated enzyme" refers to an enzyme that has been modified, for example by mutagenesis, wherein the extent of feedback inhibition of the catalytic activity of the enzyme by a metabolite is reduced such that the

enzyme exhibits enhanced activity in the presence of said metabolite compared to the unmodified enzyme. Some organisms possess deregulated forms of such enzymes as the naturally occurring, wild-type form.

The term "DNA encoding" or "encoding DNA" refers to chromosomal DNA, plasmid DNA, cDNA, plastid DNA, or synthetic DNA which codes for expression for any of the enzymes discussed herein.

The term "genome" as it applies to bacteria encompasses both the chromosome and plasmids within a bacterial host cell. Unless specified, the term "genome" as it applies to plant cells encompasses not only chromosomal or nuclear DNA found within the nucleus, but organellar DNA found within subcellular components of the cell. DNAs of the present invention introduced into plant cells can therefore be either chromosomally-integrated or organelle-localized, unless specified (e.g. "plastid genome").

The term "mutein" refers to a mutant form of a peptide, polypeptide, or protein.

"N-terminal" refers to the region of a peptide, polypeptide, or protein chain from the amino acid having a free α-amino group to the middle of the chain.

"Operably linked" refers to two amino acid or nucleotide sequences wherein one of the sequences operates to affect a characteristic of the other sequence. In the case of nucleotide sequences, for example, a promoter "operably linked" to a structural coding sequence acts to drive expression of the latter.

"Overexpression" refers to the expression of a polypeptide or protein encoded by a DNA introduced into a

host cell, wherein said polypeptide or protein is either not normally present in the host cell, or wherein said polypeptide or protein is present in said host cell at a higher level than that normally expressed from the endogenous gene encoding said polypeptide or protein.

The term "plastid" refers to the class of plant cell organelles that includes amyloplasts, chloroplasts, chromoplasts, elaioplasts, eoplasts, etioplasts, leucoplasts, and proplastids. These organelles are self-replicating, and contain what is commonly referred to as the chloroplast genome, a circular DNA molecule that ranges in size from about 120 to about 217 kb, depending upon the plant species, and which usually contains an inverted repeat region (Fosket, 1994).

The term "polyhydroxyalkanoate (PHA) synthase" refers to enzymes that convert β-hydroxyacyl-CoAs to polyhydroxy-
alkanoates and free CoA.

"Targeting sequence" refers to a nucleotide sequence which, when expressed (forming a "targeting peptide"), directs the export of an attached polypeptide to a particular cellular location, such as the chloroplast (e.g. "chloroplast targeting sequence"). The words "signal" or "transit" are equivalent to "targeting" in this context.

Production of Transgenic Plants Capable of Producing P(3HB-co-3HV) Copolymer

PHA synthesis in plants can be optimized in accordance with the present invention by expressing DNAs encoding β-ketothiolase, β-acyl-CoA reductase, and PHA

synthase in conjunction with various portions and combinations of precursor-producing enzymes, including the sequences encoding portions of the plastid PDC and the BCOADC provided herein, as discussed in the Examples below.

Plant Vectors

In plants, transformation vectors capable of introducing encoding DNAs involved in PHA biosynthesis are easily designed, and generally contain one or more DNA coding sequences of interest under the transcriptional control of 5' and 3' regulatory sequences. Such vectors generally comprise, operatively linked in sequence in the 5' to 3' direction, a promoter sequence that directs the transcription of a downstream heterologous structural DNA in a plant; optionally, a 5' non-translated leader sequence; a nucleotide sequence that encodes a protein of interest; and a 3' non-translated region that encodes a polyadenylation signal which functions in plant cells to cause the termination of transcription and the addition of polyadenylate nucleotides to the 3' end of the mRNA encoding said protein. Plant transformation vectors also generally contain a selectable marker. Typical 5'-3' regulatory sequences include a transcription initiation start site, a ribosome binding site, an RNA processing signal, a transcription termination site, and/or a polyadenylation signal. Vectors for plant transformation have been reviewed in Rodriguez et al. (1988), Glick et al. (1993), and Croy (1993).

Plant Promoters

Plant promoter sequences can be constitutive or inducible, environmentally- or developmentally-regulated, or cell- or tissue-specific. Often-used constitutive promoters include the CaMV 35S promoter (Odell et al., 1985), the enhanced CaMV 35S promoter, the Figwort Mosaic Virus (FMV) promoter (Richins et al., 1987), the mannopine synthase (*mas*) promoter, the nopaline synthase (*nos*) promoter, and the octopine synthase (*ocs*) promoter. Useful inducible promoters include heat-shock promoters (Ou-Lee et al., 1986; Ainley et al., 1990), a nitrate-inducible promoter derived from the spinach nitrite reductase gene (Back et al., 1991), hormone-inducible promoters (Yamaguchi-Shinozaki et al., 1990; Kares et al., 1990), and light-inducible promoters associated with the small subunit of RuBP carboxylase and LHCP gene families (Kuhlemeier et al., 1989; Feinbaum et al., 1991; Weisshaar et al., 1991; Lam and Chua, 1990; Castresana et al., 1988; Schulze-Lefert et al., 1989). Examples of useful tissue-specific, developmentally-regulated promoters include the β-conglycinin 7S promoter (Doyle et al., 1986; Slighton and Beachy, 1987), and seed-specific promoters (Knutzon et al., 1992; Bustos et al., 1991; Lam and Chua, 1991; Stayton et al., 1991). Plant functional promoters useful for preferential expression in seed plastids include those from plant storage protein genes and from genes involved in fatty acid biosynthesis in oilseeds. Examples of such promoters include the 5' regulatory regions from such genes as napin (Kridl et al., 1991), phaseolin, zein, soybean trypsin inhibitor, ACP,

stearoyl-ACP desaturase, and oleosin. Seed-specific gene regulation is discussed in EP 0 255 378. Promoter hybrids can also be constructed to enhance transcriptional activity (Hoffman, U.S. Patent No. 5,106,739), or to combine desired transcriptional activity and tissue specificity.

A factor to be considered in the choice of promoters is the timing of availability of the necessary substrates during expression of the PHA biosynthetic enzymes. For example, if P(3HB-co-3HV) copolymer is produced in seeds from threonine, the timing of threonine biosynthesis and the amount of free threonine are important considerations. Karchi et al. (1994) have reported that threonine biosynthesis occurs rather late in seed development, similar to the timing of seed storage protein accumulation. For example, if enzymes involved in P(3HB-co-3HV) copolymer biosynthesis are expressed from the 7S seed-specific promoter, the timing of expression thereof will be concurrent with threonine accumulation.

Plant Transformation and Regeneration

A variety of different methods can be employed to introduce such vectors into plant protoplasts, cells, callus tissue, leaf discs, meristems, etc., to generate transgenic plants, including *Agrobacterium*-mediated transformation, particle gun delivery, microinjection, electroporation, polyethylene glycol-mediated protoplast transformation, liposome-mediated transformation, etc. (reviewed in Potrykus, 1991). In general, transgenic plants comprising cells containing and expressing DNAs

encoding enzymes facilitating PHA biosynthesis can be produced by transforming plant cells with a DNA construct as described above via any of the foregoing methods; selecting plant cells that have been transformed on a selective medium; regenerating plant cells that have been transformed to produce differentiated plants; and selecting a transformed plant which expresses the enzyme-encoding nucleotide sequence.

Constitutive overexpression of, for example, a deregulated threonine deaminase employing the CaMV 35S or FMV promoter might potentially starve plants of certain amino acids, especially those of the aspartate family. If such starvation occurs, the negative effects may be avoided by supplementing the growth and cultivation media employed in the transformation and regeneration procedures with appropriate amino acids. By supplementing the transformation/regeneration media with aspartate family amino acids (aspartate, threonine, lysine, and methionine), the uptake of these amino acids into the plant can reduce any potential starvation effect caused by an overexpressed threonine deaminase. Supplementation of the media with such amino acids might thereby prevent any negative selection, and therefore any adverse effect on transformation frequency, due to the overexpression of a deregulated threonine deaminase in the transformed plant.

The encoding DNAs can be introduced either in a single transformation event (all necessary DNAs present on the same vector), a co-transformation event (all necessary DNAs present on separate vectors that are introduced into plants or plant cells simultaneously), or

by independent transformation events (all necessary DNAs present on separate vectors that are introduced into plants or plant cells independently). Traditional breeding methods can subsequently be used to incorporate the entire pathway into a single plant. Successful production of the PHA polyhydroxybutyrate in cells of *Arabidopsis* has been demonstrated by Poirier et al. (1992), and in plastids thereof by Nawrath et al. (1994).

Specific methods for transforming a wide variety of dicots and obtaining transgenic plants are well documented in the literature (Gasser and Fraley, 1989; Fisk and Dandekar, 1993; Christou, 1994; and the references cited therein).

Successful transformation and plant regeneration have been achieved in the monocots as follows: asparagus (*Asparagus officinalis*; Bytebier et al. 1987); barley (*Hordeum vulgarae*; Wan and Lemaux 1994); maize (*Zea mays*; Rhodes et al., 1988; Gordon-Kamm et al., 1990; Fromm et al., 1990; Koziel et al., 1993); oats (*Avena sativa*; Somers et al., 1992); orchardgrass (*Dactylis glomerata*; Horn et al., 1988); rice (*Oryza sativa*, including indica and japonica varieties; Toriyama et al., 1988; Zhang et al., 1988; Luo and Wu 1988; Zhang and Wu 1988; Christou et al., 1991); rye (*Secale cereale*; De la Pena et al., 1987); sorghum (*Sorghum bicolor*; Cassas et al. 1993); sugar cane (*Saccharum* spp.; Bower and Birch 1992); tall fescue (*Festuca arundinacea*; Wang et al. 1992); turfgrass (*Agrostis palustris*; Zhong et al., 1993); and wheat (*Triticum aestivum*; Vasil et al. 1992; Weeks et al. 1993; Becker et al. 1994).

Host Plants

Particularly useful plants for PHA copolymer production include those that produce carbon substrates which can be employed for PHA biosynthesis, including tobacco, wheat, potato, *Arabidopsis*, and high oil seed plants such as corn, soybean, canola, oil seed rape, sunflower, flax, and peanut. Polymers that can be produced in this manner include copolymers incorporating both short chain length and medium chain length monomers, such as P(3HB-co-3HV) copolymer.

If the host plant of choice does not produce the requisite fatty acid substrates in sufficient quantities, it can be modified, for example by mutagenesis or genetic transformation, to block or modulate the glycerol ester and fatty acid biosynthesis or degradation pathways so that it accumulates the appropriate substrates for PHA production.

Plastid Targeting of Expressed Enzymes for PHA Biosynthesis

PHA polymer can be produced in plants either by expression of the appropriate enzymes in the cytoplasm (Poirier et al., 1992) by the methods described above, or more preferably, in plastids, where higher levels of PHA production can be achieved (Nawrath et al., 1994). As demonstrated by the latter group, targeting of β-ketothiolase, acetoacetyl-CoA reductase, and PHB synthase to plastids of *Arabidopsis thaliana* results in

the accumulation of high levels of PHB in the plastids without any readily apparent deleterious effects on plant growth and seed production. As branched-chain amino acid biosynthesis occurs in plant plastids (Bryan, 1980; Galili, 1995), overexpression therein of plastid-targeted enzymes, including a deregulated form of threonine deaminase, is expected to facilitate the production of elevated levels of 2-oxobutyrate and propionyl-CoA. The latter can be condensed with acetyl-CoA by β-ketothiolase to form 3-ketovaleryl-CoA, which can then be further metabolized by a β-keto-acyl-CoA reductase to 3-hydroxyvaleryl-CoA, the precursor of the C5 subunit of P(3HB-co-3HV) copolymer. As there is a high carbon flux through acetyl-CoA in plastids, especially in seeds of oil-accumulating plants such as oilseed rape *(Brassica napus)*, canola *(Brassica rapa, Brassica campestris, Brassica carinata,* and *Brassica juncea)*, soybean (*Glycine max)*, flax (*Linum usitatissimum)*, and sunflower (*Helianthus annuus*) for example, targeting of the gene products of desired encoding DNAs to leucoplasts of seeds, or transformation of seed leucoplasts and expression therein of these encoding DNAs, are attractive strategies for achieving high levels of PHA biosynthesis in plants.

All of the enzymes discussed herein can be modified for plastid targeting by employing plant cell nuclear transformation constructs wherein DNA coding sequences of interest are fused to any of the available transit peptide sequences capable of facilitating transport of the encoded enzymes into plant plastids

(partially summarized in von Heijne et al., 1991), and driving expression by employing an appropriate promoter. The sequences that encode a transit peptide region can be obtained, for example, from plant nuclear-encoded plastid proteins, such as the small subunit (SSU) of ribulose bisphosphate carboxylase, plant fatty acid biosynthesis related genes including acyl carrier protein (ACP), stearoyl-ACP desaturase, β-ketoacyl-ACP synthase and acyl-ACP thioesterase, or LHCPII genes. The encoding sequence for a transit peptide effective in transport to plastids can include all or a portion of the encoding sequence for a particular transit peptide, and may also contain portions of the mature protein encoding sequence associated with a particular transit peptide. Numerous examples of transit peptides that can be used to deliver target proteins into plastids exist, and the particular transit peptide encoding sequences useful in the present invention are not critical as long as delivery into a plastid is obtained. Proteolytic processing within the plastid then produces the mature enzyme. This technique has proven successful not only with enzymes involved in PHA synthesis (Nawrath et al., 1994), but also with neomycin phosphotransferase II (NPT-II) and CP4 EPSPS (Padgette et al., 1995), for example.

Of particular interest are transit peptide sequences derived from enzymes known to be imported into the leucoplasts of seeds. Examples of enzymes containing useful transit peptides include those related to lipid biosynthesis (e.g., subunits of the plastid-targeted dicot acetyl-CoA carboxylase, biotin

carboxylase, biotin carboxyl carrier protein, α-carboxytransferase, plastid-targeted monocot multifunctional acetyl-CoA carboxylase (Mr, 220,000); plastidic subunits of the fatty acid synthase complex (e.g., acyl carrier protein (ACP), malonyl-ACP synthase, KASI, KASII, KASIII, etc.); steroyl-ACP desaturase; thioesterases (specific for short, medium, and long chain acyl ACP); plastid-targeted acyl transferases (e.g., glycerol-3-phosphate: acyl transferase); enzymes involved in the biosynthesis of aspartate family amino acids; phytoene synthase; gibberellic acid biosynthesis (e.g., *ent*-kaurene synthases 1 and 2); sterol biosynthesis (e.g., hydroxy methyl glutaryl-coA reductase); and carotenoid biosynthesis (e.g., lycopene synthase).

Exact translational fusions to the transit peptide of interest may not be optimal for protein import into the plastid. By creating translational fusions of any of the enzymes discussed herein to the precursor form of a naturally imported protein or C-terminal deletions thereof, one would expect that such translational fusions would aid in the uptake of the engineered precursor protein into the plastid. For example, Nawrath et al., (1994) used a similar approach to create the vectors employed to introduce the PHB biosynthesis genes of *A. eutrophus* into *Arabidopsis*.

It is therefore fully expected that targeting of the enzymes discussed herein to leaf chloroplasts or seed plastids such as leucoplasts by fusing transit peptide gene sequences thereto will further enhance *in vivo* conditions for the biosynthesis of PHAs,

especially P(3HB-co-3HV) copolymer, in plants.

Plastid Transformation for Expression of Enzymes Involved in PHA Biosynthesis

Alternatively, enzymes facilitating the biosynthesis of metabolites such as threonine, 2-oxobutyrate, propionyl-CoA, 3-ketovaleryl-CoA, 3-hydroxy-valeryl-CoA, and PHAs discussed herein can be expressed *in situ* in plastids by direct transformation of these organelles with appropriate recombinant expression constructs. Constructs and methods for stably transforming plastids of higher plants are well known in the art (Svab et al., 1990; Svab et al., 1993; Staub et al., 1993; Maliga et al., U.S. Patent No. 5,451,513; PCT International Publications WO 95/16783, WO 95/24492, and WO 95/24493). These methods generally rely on particle gun delivery of DNA containing a selectable marker in addition to introduced DNA sequences for expression, and targeting of the DNA to the plastid genome through homologous recombination. Transformation of a wide variety of different monocots and dicots by particle gun bombardment is routine in the art (Hinchee et al., 1994; Walden and Wingender, 1995).

DNA constructs for plastid transformation generally comprise a targeting segement comprising flanking DNA sequences substantially homologous to a predetermined sequence of a plastid genome, which targeting segment enables insertion of DNA coding sequences of interest into the plastid genome by homologous recombination with said predetermined

sequence; a selectable marker sequence, such as a sequence encoding a form of plastid 16S ribosomal RNA that is resistant to spectinomycin or streptomycin, or that encodes a protein which inactivates spectinomycin or streptomycin (such as the *aadA* gene), disposed within said targeting segment, wherein said selectable marker sequence confers a selectable phenotype upon plant cells, substantially all the plastids of which have been transformed with said DNA construct; and one or more DNA coding sequences of interest disposed within said targeting segment relative to said selectable marker sequence so as not to interfere with conferring of said selectable phenotype. In addition, plastid expression constructs also generally include a plastid promoter region and a transcription termination region capable of terminating transcription in a plant plastid, wherein said regions are operatively linked to the DNA coding sequences of interest.

A further refinement in chloroplast transformation/expression technology that facilitates control over the timing and tissue pattern of expression of introduced DNA coding sequences in plant plastid genomes has been described in PCT International Publication WO 95/16783. This method involves the introduction into plant cells of constructs for nuclear transformation that provide for the expression of a viral single subunit RNA polymerase and targeting of this polymerase into the plastids via fusion to a plastid transit peptide. Transformation of plastids with DNA constructs comprising a viral single subunit RNA polymerase-specific promoter specific to the RNA polymerase expressed from the nuclear expression

constructs operably linked to DNA coding sequences of interest permits control of the plastid expression constructs in a tissue and/or developmental specific manner in plants comprising both the nuclear polymerase construct and the plastid expression constructs. Expression of the nuclear RNA polymerase coding sequence can be placed under the control of either a constitutive promoter, or a tissue- or developmental stage-specific promoter, thereby extending this control to the plastid expression construct responsive to the plastid-targeted, nuclear-encoded viral RNA polymerase. The introduced DNA coding sequence can be a single encoding region, or may contain a number of consecutive encoding sequences to be expressed as an engineered or synthetic operon. The latter is especially attractive where, as in the present invention, it is desired to introduce multigene biochemical pathways into plastids. This approach is not practical using standard nuclear transformation techniques since each gene introduced therein must be engineered as a monocistron, including an encoded transit peptide and appropriate promoter and terminator signals. Individual gene expression levels may vary widely among different cistrons, thereby possibly adversely affecting the overall biosynthetic process. This can be avoided by the chloroplast transformation approach.

Production of Transgenic Plants Comprising Genes for PHA Biosynthesis

Plant transformation vectors capable of delivering DNAs (genomic DNAs, plasmid DNAs, cDNAs, or synthetic DNAs) encoding PHA biosynthetic enzymes and other enzymes for optimizing substrate pools for PHA biosynthesis as discussed in Examples 1-7 herein can be easily designed. Various strategies can be employed to introduce these encoding DNAs to produce transgenic plants capable of biosynthesizing high levels of PHAs, including:

1. Transforming individual plants with an encoding DNA of interest. Two or more transgenic plants, each containing one of these DNAs, can then be grown and cross-pollinated so as to produce hybrid plants containing the two DNAs. The hybrid can then be crossed with the remaining transgenic plants in order to obtain a hybrid plant containing all DNAs of interest within its genome.

2. Sequentially transforming plants with plasmids containing each of the encoding DNAs of interest, respectively.

3. Simultaneously cotransforming plants with plasmids containing each of the encoding DNAs, respectively.

4. Transforming plants with a single plasmid containing two or more encoding DNAs of interest.

5. Transforming plants by a combination of any of the foregoing techniques in order to obtain a plant that expresses a desired combination of encoding DNAs of interest.

Traditional breeding of transformed plants produced according to any one of the foregoing methods by successive rounds of crossing can then be carried out to incorporate all the desired encoding DNAs in a single homozygous plant line (Nawrath et al., 1994; PCT International Publication WO 93/02187). Similar strategies can be employed to produce bacterial host cells engineered for optimal PHA production.

In methods 2 and 3, the use of vectors containing different selectable marker genes to facilitate selection of plants containing two or more different encoding DNAs is advantageous. Examples of useful selectable marker genes include those conferring resistance to kanamycin, hygromycin, sulphonamides, glyphosate, bialaphos, and phosphinothricin.

Stability of Transgene Expression

As several overexpressed enzymes may be required to produce optimal levels of substrates for copolymer formation, the phenomenon of co-suppression may influence transgene expression in transformed plants. Several strategies can be employed to avoid this potential problem (Finnegan and McElroy, 1994).

One commonly employed approach is to select and/or screen for transgenic plants that contain a single intact copy of the transgene or other encoding DNA

(Assaad et al., 1993; Vaucheret, 1993; McElroy and Brettell, 1994). *Agrobacterium*-mediated transformation technologies are preferred in this regard.

Inclusion of nuclear scaffold or matrix attachment regions (MAR) flanking a transgene has been shown to increase the level and reduce the variability associated with transgene expression in plants (Stief et al., 1989; Breyne et al., 1992; Allen et al., 1993; Mlynarova et al., 1994; Spiker and Thompson, 1996). Flanking a transgene or other encoding DNA with MAR elements may overcome problems associated with differential base composition between such transgenes or encoding DNAs and integrations sites, and/or the detrimental effects of sequences adjacent to transgene integration sites.

The use of enhancers from tissue-specific or developmentally-regulated genes may ensure that expression of a linked transgene or other encoding DNA occurs in the appropriately regulated manner.

The use of different combinations of promoters, plastid targeting sequences, and selectable markers for introduced transgenes or other encoding DNAs can avoid potential problems due to *trans*-inactivation in cases where pyramiding of different transgenes within a single plant is desired.

Finally, inactivation by co-suppression can be avoided by screening a number of independent transgenic plants to identify those that consistently overexpress particular introduced encoding DNAs (Register et al., 1994). Site-specific recombination in which the endogenous copy of a gene is replaced by the same gene, but with altered expression characteristics, should

obviate this problem (Yoder and Goldsbrough, 1994).

Any of the foregoing methods, alone or in combination, can be employed in order to insure the stability of transgene expression in transgenic plants of the present invention.

Cloning of plastid pyruvate dehydrogenase complex and branched chain oxoacid dehydrogenase complex subunits and components

The present invention provides nucleotide sequences that encode the E1α and E1β subunits, and the E2 component, of the plastid pyruvate dehydrogenase complex, as well as the E1α and E1β subunits, and the E2 component, of the branched chain oxoacid dehydrogenase complex, of *Arabidopsis thaliana*. These sequences can be cloned by any appropriate method known in the art. For example, cDNA clones of known components of similar enzymes from other species can be utilized to screen a cDNA library from which the cDNA for the enzyme component is desired. Sources from which the plastid PDC E1α and E1β cDNAs can be obtained include the analogous enzyme-encoding cDNAs from the red alga *Porphyra purpurea*; for the E2 component of the plastid pyruvate dehydrogenase, the analogous enzyme gene from the cyanobacterium *Synechocystis* sp. can be used. The cDNA for the E1α of a BCOADC can be isolated by identifying cDNAs which have significant homology to analogous tomato, human and bovine BCOADC E1α sequences. Similarly, the E1β and the E2 components of a BCOADC can be isolated by comparing the similarity of candidate sequences with the human and bovine BCOADC

E1β and E2 components, respectively. A cDNA library for the isolation of these components can be an expressed sequence tag library, for example one comprising cDNA from *Arabidopsis thaliana*.

The cloned cDNAs for the plastid PDC and the BCOADC components can be sequenced in order to determine the nucleotide sequence and deduce the amino acid sequence for these enzymes. The sequences of these cDNAs can be determined by any method known in the art. Methods for the determination of various portions of the sequenced cDNA, such as a plastid targeting sequence, are also well known in the art.

Engineering plants to produce propionyl-CoA in plastids

The production of the P(3HB-co-3HV) precursor propionyl-CoA in plastids requires the presence of two elements which are not present, or which are present at very low levels, in the plastids of wild-type plants: 2-oxobutyrate, and enzymes which will convert 2-oxobutyrate into propionyl-CoA.

As noted above, Gruys et al. (1998) discusses several methods for the production of 2-oxobutyrate in plastids. These include:

--Overexpression of threonine deaminase;

--Overexpression of aspartate kinase and threonine deaminase; and

--Overexpression of aspartate kinase, homoserine dehydrogenase, and threonine deaminase.

The overexpression of these enzymes can be accomplished through the transformation into plants of nucleotide sequences encoding these enzymes, operably linked to a plant promoter, such as the cauliflower

mosaic virus (CaMV) 35s promoter, or any other promoter known in the art which causes overexpression of such enzymes in plants.

The expression of these and other enzymes in plastids can be achieved in at least two ways:

1. By transforming coding sequences for these enzymes directly into the plastid genome in such a way that they are incorporated into the plastid genome. Constructs and methods for stably transforming plastids of higher plants are well known in the art (for example, Svab et al., 1990; Svab et al., 1993; Staub et al., 1993; Maliga et al., U.S. Patent No. 5,451,513; PCT International Publications WO 95/16783, WO 95/24492, and WO 95/24493). These methods generally rely on particle gun delivery of DNA containing a selectable marker in addition to introduced DNA sequences for expression, and targeting of the DNA to the plastid genome through homologous recombination.

2. By creating a plant transformation vector comprising a coding sequence for the enzyme operably linked to a plastid targeting sequence, then transforming this vector into the plant. All of the enzymes discussed herein can be modified for plastid targeting by employing plant cell nuclear transformation constructs wherein DNA coding sequences of interest are fused to any of the available targeting peptide sequences capable of facilitating transport of the encoded enzymes into plant plastids, and driving expression by employing an appropriate promoter. Examples of plastid targeting peptides are provided in

Table 1 and in von Heijne et al. (1991). The sequences that encode a targeting peptide region can be obtained, for example, from plant nuclear-encoded plastid proteins, such as the small subunit (SSU) of ribulose bisphosphate carboxylase, plant fatty acid biosynthesis related genes including acyl carrier protein (ACP), stearoyl-ACP desaturase, β-ketoacyl-ACP synthase and acyl-ACP thioesterase, or LHCPII genes. The encoding sequence for a targeting peptide effective in transport to plastids can include all or a portion of the encoding sequence for a particular targeting peptide, and can also contain portions of the mature protein encoding sequence associated with a particular targeting peptide. Numerous examples of targeting peptides that can be used to deliver target proteins into plastids exist, and the particular targeting peptide encoding sequences useful in the present invention are not critical as long as delivery into a plastid is obtained. Proteolytic processing within the plastid then produces the mature enzyme. This technique has proven successful not only with enzymes involved in PHA synthesis (Nawrath et al., 1994), but also with neomycin phosphotransferase II (NPT-II) and CP4 EPSPS (Padgette et al., 1995), for example.

Table 1. Examples of plastid proteins from various species with known plastid targeting sequences that can be used to target proteins to plastids

Chloroplast Targeting Peptides

Arabidopsis thaliana:

5-enolpyruvyl-shikimate-3-phosphate synthase
Rubisco activase
Rubisco small subunit
Tryptophan synthase

Brassica napus:

Acyl carrier protein
Plastid chaperonin-60

Pisum sativum:

Carbonic anhydrase
Chloroplast stromal HSP70
Glutamine synthetase
Rubisco small subunit

Reference: von Heijne, G.; Hirai, T.; Klosgen, R.B.; Steppuhn, J.; Bruce, B.; Keegstra, K.; Herrmann, R. (1991) CHLPEP-A database of chloroplast transit peptides. *Plant Molecular Biology Reporter* 9:104-126.

Engineering plants to produce poly(3-hydroxybutyrate-3-hydroxyvalerate) copolymer

Plants which produce P(3HB-co-3HV) can be created by engineering them to produce 2-oxobutyrate, to convert 2-oxobutyrate to propionyl-CoA, and to synthesize P(3HB-co-3HV) from propionyl-CoA and acetyl-CoA. Methods for producing plants which synthesize 2-oxobutyrate are discussed above. Such plants can be modified to convert 2-oxobutyrate to propionyl-CoA in the manner discussed below.

The nucleotide sequences of the BCOADC E1α and E1β subunits, and that of the E2 component, are provided herein as a means to effect the conversion of 2-oxobutyrate to propionyl-CoA in plastids containing the 2-oxobutyrate substrate. It is not necessary to provide the E3 component since the E3 components of all of the α-ketoacid dehydrogenase complexes are probably interchangeable. The E3 subunit already present in the plastid PDC thus almost certainly functions with plastid-targeted BCOADC subunits. The nucleotide sequences of the plastid PDC E1α and E1β subunits, and the E2 component, provide sources of plastid targeting sequences. These plastid PDC sequences can also be genetically manipulated to enhance their ability to convert 2-oxobutyrate to propionyl-CoA, as suggested by Gruys et al. (1998).

The nucleotide sequences encoding the BCOADC E1α and E1β subunits, and the E2 component, can be directly

transformed into the plastid genome by the methods discussed above. Alternatively, the BCOADC E1 and E2 nucleotide sequences can be transformed into the plant nuclear genome, wherein the enzyme coding sequences are operably linked to a plastid targeting sequence by methods known in the art. See Example 7. Useful plastid targeting sequences include those from the plastid PDC. These targeting sequences from *Arabidopsis thaliana* are disclosed in Examples 1 and 2, below.

As another alternative for utilizing a BCOADC for the conversion of 2-oxobutyrate to propionyl-CoA in plastids, a nucleotide sequence encoding the BCOADC E1β subunit can be engineered to utilize the PDC E2 component which is already present in the plastids. The BCOADC E1β subunit can be modified such that the native E2 binding region thereof is replaced with the E2 binding region of the plastid PDC E1β subunit. The nucleotide sequences encoding the modified BCOADC E1β subunit and the BCOADC E1α subunit can also be operably linked to a plastid targeting sequence. The modified nucleotide sequences for these two subunits (α and β) of the BCOADC E1 component can then be inserted into plants by standard plant transformation methods, where they are translated in the cytoplasm. The enzymes are then transported to the plastid where they combine with the plastid PDC E2 and E3 components, and catalyze the conversion of 2-oxobutyrate to propionyl-CoA. See Example 6 below.

The conversion of propionyl-CoA and acetyl-CoA to P(3HB-co-3HV) requires a β-ketothiolase, a β-ketoacyl-CoA reductase, and a PHA synthase. Nucleotide

sequences encoding these enzymes can be incorporated into the plastid genome directly, or into the nuclear genome, with operably linked plastid targeting sequences, utilizing the same well-known methods as previously discussed. Preferred β-ketothiolases are BktB and pAE65 from *A. eutrophus*, and *Zoogloea ramigera* β-ketothiolases "A" and "B", as disclosed in Gruys et al (1998). Preferred β-ketoacyl-CoA reductases and PHA synthases include those from *A. eutrophus*, encoded by the phbB and phbC genes, respectively. However, the use of other β-ketothiolases which are able to utilize propionyl-CoA, and the use of other β-ketoacyl-CoA reductases and PHA synthases are within the scope of this invention. Included are those enzymes from, for example, *Alcaligenes faecalis*, *Aphanothece* sp., *Azotobacter vinelandii*, *Bacillus cereus*, *Bacillus megaterium*, *Beijerinkia indica*, *Derxia gummosa*, *Methylobacterium* sp., *Microcoleus* sp., *Nocardia corallina*, *Pseudomonas cepacia*, *Pseudomonas extorquens*, *Pseudomonas oleovorans*, *Rhodobacter sphaeroides*, *Rhodobacter capsulatus*, *Rhodospirillum rubrum*, and *Thiocapsa pfennigii*.

P(3HB-co-3HV) Copolymer Composition

The P(3HB-co-3HV) copolymers of the present invention can comprise about 75-99% 3HB and about 1-25% 3HV based on the total weight of the polymer. More preferably, P(3HB-co-3HV) copolymers of the present invention comprise about 85-99% 3HB and about 1-15% 3HV. Even more preferably, such copolymers comprise about 90-99% 3HB and about 1-10% 3HV. P(3HB-co-3HV) copolymers comprising about 4%, about 8%, and about 12% 3HV on a weight basis possess properties that have made

them commercially attractive for particular applications. One skilled in the art can modify P(3HB-co-3HV) copolymers of the present invention by physical or chemical means to produce copolymer derivatives having desirable properties different from those of the plant-produced copolymer.

Optimization of P(3HB-co-3HV) copolymer production by the methods discussed herein is expected to result in yields of copolymer in the range of from at least about 1% to at least about 20% of the fresh weight of the plant tissue, organ, or structure in which it is produced.

The following examples illustrate the invention, but are not to be taken as limiting the various aspects of the invention so illustrated.

Conventional methods of gene isolation, molecular cloning, vector construction, etc., are well known in the art and are summarized in Sambrook et al., 1989, and Ausubel et al., 1989 and 1994. One skilled in the art can readily repeat the methods and reproduce the compositions described herein without undue experimentation. The various DNA sequences, fragments, etc., necessary for this purpose can be readily obtained as components of commercially available plasmids, or synthesized by well known methods, or are otherwise well known in the art and publicly available.

Example 1

Cloning and Sequencing cDNA Encoding the E1α and E1β Subunits of the *Arabidopsis thaliana* Plastid Pyruvate Dehydrogenase Complex

Expressed sequence tag (EST) clones (Reith et al., 1995) from the Arabidopsis Biological Resource Center (ABRC) at Ohio State University were used to isolate full-length cDNAs for both the plastid E1α and E1β subunits from an *A. thaliana* cDNA library. Two clones (GenBank accessions T75600 and N65566) were identified as potentially encoding the plastid E1α and E1β subunits as follows.

Oligonucleotides were designed based on sequences common to *P. purpurea odpA* and *odpB* and the two *Arabidopsis* EST sequences and synthesized (all recited in the 5'-3' direction):

E1α: 5′ primer, CGGTACtCAAGTCTGACTCTGTCGTT (SEQ ID NO:7);

3′ primer, CCTTCGAuAGGTTCCATCTCCGAAAAA (SEQ ID NO:8);

E1β: 5′ primer, CGGTACtCTTCGAGGCTCTTCAGGAA (SEQ ID NO:9);

3′ primer, CCTTCGAuACGGGCCTTAGACCAGT (SEQ ID NO:10).

The symbols denote restriction sites (t: *Kpn* I, and u: *Hind* III) added for subcloning. Thermal cycling was used to amplify cDNA fragments from *A. thaliana* using first strand cDNA. Thermal cycling reactions (50 μl total volume) contained 10 mM Tris-HCl, pH 7.9, 1.25 mM $MgCl_2$, 25 μM dNTPs, 5 units *Taq* polymerase (Promega, Madison, WI), 2 μg *A. thaliana* first strand cDNA, and 10 ng of each primer. Thermal cycling was performed with a Perkin-Elmer model 480, with rapid ramp times set at 1°C/s. Cycling conditions were 94°C for 20 s, 50°C for 30 s, 72°C for 2 min with 6 s extensions each cycle and 30 rounds of cycling. Under these conditions, products containing 288 base pairs (E1α)

and 215 base pairs (E1β) were obtained. The products were subcloned into pGEMT (Promega, Madison, WI) and sequenced to confirm their identity. Thermal cycling was also used to generate probes radiolabelled with (α^{32}P)-dCTP, using reaction mixtures identical to those previously described except for a 1000-fold reduction in the concentration of non-radioactive dCTP. Before use, the probes were desalted using Sephadex G-50 columns to remove unincorporated nucleotides. An *Arabidopsis* cDNA library (λ-PRL2, obtained from the ABRC) was plated at a density of $2.25x10^4$ plaques per plate for a total of $2.25x10^5$ plaques. BioTrace NT nylon filters (Gelman, Ann Arbor, MI) were used for plaque-lifts and were processed according to the manufacturer's specifications. Hybridizations were performed according to *Current Protocols in Molecular Biology* (Ausubel et al., 1994). After three rounds of screening, 7 potential E1α and 12 potential E1β cDNA clones were isolated, ranging in size from 1100 to 1550 base pairs. Plaque-purified λ phage were treated according to the manufacturer's instructions (Gibco BRL, Gaithersburg, MD) in order to excise the pZL-1 recombinant clones.

DNA sequencing was performed using an ABI prism Model 377 sequencer, and analyzed using IntelliGenetics GeneWorks DNA analysis program version 2.5 on a Macintosh computer. Dye-deoxy terminating cycle sequencing reactions were carried out on both strands of full-length cDNA inserts and deletion fragments derived therefrom.

DNA isolation and Northern and Southern blotting

were carried out according to *Current Protocols in Molecular Biology* (Sections 2.9.1, 4.3.1 and 4.9.1; Ausubel et al., 1994). RNA isolation was accomplished with the RNAgents total RNA isolation kit (Promega, Madison, WI). Northern blot prehybridization (3 h), hybridization (12 h), and 4 washes were done with 2.5 X SSPE (1X = 0.15 mM NaCl, 0.02 mM Na_2PO_4, 2 μM EDTA, pH 7.4), 1% SDS, 1% non-fat dry milk, and 250 μg/ml salmon sperm DNA at 68°C. Blots were exposed on Kodak X-OMAT/AR film (Rochester, New York) at -70°C with an intensifying screen.

Among the genes present in the *P. purpurea* plastome are two open reading frames, *odpA* and *odpB*, encoding proteins 32% identical to the *Arabidopsis* mitochondrial E1α and E1β subunits (Grof et al., 1995; Leuthy et al., 1994; Leuthy et al., 1995). Attempts to use cloned mitochondrial PDC cDNAs as probes to identify plastid sequences have been unsuccessful. Based upon the *odpA* and *odpB* sequences, two EST clones (accessions T75600 and N65566) which appear to encode proteins more highly related to the *P. purpurea odpA* and *odpB* sequences than to the *Arabidopsis* mitochondrial sequences were used to isolate two cDNAs as potential E1α and E1β clones.

The nucleotide sequence of the *Arabidopsis* plastid PDC E1α cDNA (Genbank Accession No. U80185) is shown in Appendix A and as SEQ ID NO:1. E1α cDNA (1530 bp) has a 106 bp 5' untranslated region, a 1284 bp open reading frame encoding a polypeptide of 428 amino acids (Appendix B and SEQ ID NO:2), and a 140 bp 3' untranslated region. The nucleotide sequence of the

Arabidopsis plastid PDH E1β cDNA (Genbank Accession No. U80186) is shown in Appendix C and as SEQ ID NO:3. The E1β cDNA (1441 bp) has a 6 bp 5' untranslated region, a 1218 bp open reading frame encoding a polypeptide of 406 amino acids (Appendix D and SEQ ID NO:4), and a 217 bp 3' untranslated region. The calculated molecular weight and isoelectric point values for the E1α and E1β polypeptides encoded by the open reading frames are 47,120 with a pI of 7.25, and 44,208 with a pI of 5.89, respectively. The deduced amino acid sequence for E1α has 61%, and E1β 68%, identity with *P. purpurea odpA* and *odpB*, respectively.

The first 68 residues of E1α and the first 73 residues of E1β exhibit characteristics of chloroplast targeting peptides but not those of mitochondrial targeting sequences (Gavel et al., 1990; von Heijne et al., 1989). To determine structural motifs of the targeting peptides, we used the GeneWorks (IntelliGenetics, Mountain View, CA) protein algorithm to identify possible α-helix and β-strands. Both plastid E1α and E1β have the potential to form amphiphilic β-strands consistent with plastid targeting sequences, but did not fit the amphiphilic α-helix which is characteristic of mitochondrial targeting sequences.

~~Tables 2 and 3~~ Fig. 8 and 9 show the alignment of the deduced amino acid sequences of PDH E1α and E1β. Abbreviations are the same as in Fig 7. * indicates conserved, • non-conserved phosphorylation sites. ° indicates the conserved Cys 62 of the mature *H.s.* E1α sequence.

Overall, there is 28% sequence identity between

Arabidopsis plastid PDC E1α and its mammalian counterparts. However, in specific regions, the degree of sequence conservation is much higher. The PDH component of PDC requires thiamine pyrophosphate (TPP) as a cofactor for decarboxylation of pyruvate (Patel et al., 1990). It has been reported that TPP binds to the E1α subunit of mammalian PDH at a site containing a structural motif common to pyrophosphate-binding enzymes (Reed, 1974). A similar motif (50% identity with the bovine E1α TPP-binding domain) is found in the *A. thaliana* plastid E1α sequence at residues 160-213 (~~Table 2~~ Fig. 9).

A highly conserved Cys residue (Cys 62 of mature human E1α, ~~Table 2~~ Fig. 9) has been identified in eukaryotic PDH E1α sequences, and it has been proposed that this Cys is an essential component of the enzyme's active site (Ali et al., 1993). The *A. thaliana* plastid E1α sequence contains a similar motif, i.e. the same immediate flanking residues at 112-116, but the otherwise conserved Cys is replaced with a Val (~~Table 2~~ Fig. 9).

Mitochondrial PDCs are regulated in part by reversible phosphorylation of three conserved Ser residues in the E1α sequence by a specific, complex-associated PDH-kinase (Reed, 1974). The Ser residues phosphorylated in mammalian mitochondrial PDH are also conserved in the plant mitochondrial (Luethy et al., 1995), yeast (Behal et al., 1989), and nematode (Johnson et al., 1992) amino acid sequences. However, while the plant mitochondria PDC is reversibly phosphorylated (Randall et al., 1989; Randall et al., 1996), all evidence to date indicates that plastid PDC

activity is not regulated by phosphorylation (Camp et al., 1985). Despite this difference, the regulatory Ser residues and their flanking sequences are present in the plastid E1α sequence (~~Table 2~~ Fig. 2). Korotchkina and Patel (1995) have reported the results from mutagenesis of these phosphorylation sites, and concluded that site one is closer to the active site or lies on the pathway to the main catalytic conformational change. This might explain why this region is so highly conserved. The amino acid-motif corresponding to phosphorylation site one in mitochondrial PDH sequences is present in the plastid polypeptide (Tyr 320-Pro 330 or Tyr 287-Pro 297 in the *H. s.* sequence, ~~Table 2~~ Fig. 9). Two of the four substitutions are by residues with conserved properties. The sequence of the plastid E1α corresponding to phosphorylation site two lacks a Ser and the region is dominated by five acidic and two basic residues (Asp 329-Asp 339). The *Arabidopsis* plastid E1α sequence contains a Ser at site 3 (Ala 259-Ala 267), but the flanking residues are dissimilar to the mammalian site 3 (~~Table 2~~ Fig. 8). While two of the three Ser are in the appropriate positions, it is most likely then that plastid PDC is not regulated by phosphorylation due to the lack of plastid PDH-kinase (Camp et al., 1985).

Wexler et al. (1991) compared alignments of three PDH and three branched-chain α-keto acid dehydrogenase sequences. Among E1β sequences, four regions of sequence conservation were observed. Region one, the proposed E2 interaction site, is present in the *Arabidopsis* plastid PDH E1β sequence (~~Table 3~~ Fig. 9). Conserved regions two and three share high homology

with other decarboxylating enzymes, suggesting a role in decarboxylation of pyruvate (Wexler et al., 1991). A functional role has not yet been attributed to region four (~~Table 3~~ Fig. 9). Eswaran et al. (1995) have described Arg 239 as being an essential residue near or at the active site of the bovine E1β. This residue is conserved throughout the eukaryotic PDH sequences (e.g., Arg 269 of *H. s.* sequence in ~~Table 3~~ Fig. 9), and is present in the *A. thaliana* plastid E1β sequence at position 318.

The genomic organization of *Arabidopsis* E1α and E1β was determined by Southern blot analysis. An E1α cDNA probe hybridized to a single restriction fragment in each lane, suggesting one gene (Fig. 4A). An E1β cDNA probe, on the other hand, hybridized to multiple fragments in a pattern consistent with the restriction digest of E1β cDNA (data not shown). The *Xba* I lane contained multiple hybridizing bands which could be due to a second gene or an intron containing an *Xba* I restriction site (Fig. 4B).

In order to evaluate expression of the *A. thaliana* plastid PDH genes, 10 μg total RNA obtained from young leaves were resolved by formaldehyde gel electrophoresis. Northern blot analyses confirmed the expression of a single mRNA species of 1.65 kb for E1α and 1.5 kb for E1β (Figs. 5A and 5B).

The two cDNAs reported here have been identified as encoding plastid rather than mitochondrial proteins based on their high homology with the *P. purpurea* chloroplast genes, the presence of N-terminal sequences characteristic of plastid targeting peptides, and their

relatively low homology with plant mitochondrial E1 subunits (Grof et al., 1995; Leuthy et al., 1994; Leuthy et al., 1995). Assessments of the mature N-terminal sequences were based on homology with the mature *odp* and mitochondrial E1 sequences.

The mature *A. thaliana* plastid E1α and E1β amino acid sequence have the highest homology (68%) with the *P. purpurea* chloroplast *odpA* and *odpB* sequences, respectively, but only 31 and 32% identity with the respective *A. thaliana* mitochondrial E1 sequences (~~Tables 2 and~~ Figs. 8A and 9). The homology with other eukaryotic mitochondrial E1 sequences is lower yet. Additionally, a monoclonal antibody prepared against mitochondrial E1α does not recognize chloroplastic E1α (Luethy et al., 1995) nor does the monoclonal antibody recognize the recombinant plastid E1α on immunoblots.

Dendrogram analyses show that *A. thaliana* plastid E1, *P. purpurea* chloroplast *odp*, and *Synechocystis sp.* (a cyanobacterium) *pdh* sequences segregate as a family distinct from mitochondrial and bacterial sequences (Figs. 6A and 6B). A similar separation has also been shown for plastid and mitochondrial ribosomal RNA sequences (Palmer, 1992). The *A. thaliana* plastid cDNAs and *P. purpurea odp* genes are the only sequences reported thus far for plastid forms of PDH.

As additional cDNAs and genes for plastid and mitochondrial specific isozymes are determined, insight as to the lineage of plastid genes will be gained. Mitochondrial rRNA genes show convincing similarity to purple-photosynthetic bacterial rRNA sequences. In contrast, plastid rRNA has similarity with

cyanobacterial rRNA. This relationship between plastids and cyanobacteria has also been noted for genes encoding the transcriptional and translational apparatus (Palmer, 1992). The new sequences reported here should contribute to understanding if the emergence of mitochondria and plastids was the result of single or multiple primary (i.e., eubacteria/eukaryotic) endosymbioses, or if secondary (i.e., eukaryotic/eukaryotic) endosymbioses led to the establishment of these organelles (Palmer, 1992).

Antibodies to the E1α subunit of the plastid pyruvate dehydrogenase complex were generated by inserting the gel purified BamHI to HindIII fragment of the cDNA for E1 at the BamHI (5') to HindIII (3') cloning site of pET28a (Novagen). The recombinant clone was expressed, and the 5' end sequenced to ensure the correct reading frame. The recombinant protein was expressed using the above construct in *E. coli* strain BL21 (DE3) (Novagen). Growth conditions were as follows: A single colony was picked and cultured in 5 mL LB + 150 micrograms ampicillin overnight at 37 C shaking at 200 rpm. The 5ml culture was used to inoculate 500 mL LB + 150 microgram ampicillin and was allowed to grow for 4 h. The culture was then induced using 0.1 mM IPTG and allowed to shake at 37 C for an additional 5 h. The culture was then centrifuged in a GSA rotor at 7,000 rpm to pellet cells. Cells were lysed in 6 M guanidinium HCl, 10 mM Tris pH 8.0 at room temperature. Cell debris was pelleted at 12,000 rpm in an SS-34 rotor for 20 min, and the recombinant protein was purified using Ni-NTA agarose. Rabbits were injected with 150 microgram of recombinant protein

mixed 1:1 with complete adjuvant. A 30 day boost was given with the same protein preparation, at the same concentration. Ten days after the boost, the antibody titer was determined to be 1:80,000 against pea chloroplast stromal extract by immunoblot procedures.

It should be noted that the present invention encompasses not only the specific DNA sequences disclosed herein and the polypeptides encoded thereby, but also biologically functional equivalent nucleotide and amino acid sequences. The phrase "biologically functional equivalent nucleotide sequences" denotes DNAs and RNAs, including chromosomal DNA, plasmid DNA, cDNA, synthetic DNA, and mRNA nucleotide sequences, that encode polypeptides exhibiting the same or similar enzymatic activity as that of the enzyme polypeptides encoded by the sequences disclosed herein when assayed by standard enzymatic methods, or by complementation. Such biologically functional equivalent nucleotide sequences can encode polypeptides that contain a region or moiety exhibiting sequence similarity to the corresponding region or moiety of the present disclosed polypeptides.

One can isolate polypeptides useful in the present invention from various organisms based on homology or sequence identity. Although particular embodiments of nucleotide sequences encoding the polypeptides disclosed herein are shown in the various SEQ IDs presented, it should be understood that other biologically functional equivalent forms of such polypeptide-encoding nucleic acids can be readily isolated using conventional DNA-DNA or DNA-RNA hybridization techniques. Thus, the present invention

also includes nucleotide sequences that hybridize to any of the nucleic acid SEQ IDs and their complementary sequences presented herein, and that code on expression for polypeptides exhibiting the same or similar enzymatic activity as that of the presently disclosed polypeptides. Such nucleotide sequences preferably hybridize to the nucleic acid sequences presented herein or their complementary sequences under moderate to high stringency (see Sambrook et al., 1989). Exemplary conditions include initial hybridization in 6X SSC, 5X Denhardt's solution, 100 μg/ml fish sperm DNA, 0.1% SDS, at 55°C for sufficient time to permit hybridization (e.g., several hours to overnight), followed by washing two times for 15 min each in 2X SSC, 0.1% SDS, at room temperature, and two times for 15 min each in 0.5-1X SSC, 0.1% SDS, at 55°C, followed by autoradiography. Typically, the nucleic acid molecule is capable of hybridizing when the hybridization mixture is washed at least one time in 0.1X SSC at 55°C, preferably at 60°C, and more preferably at 65°C.

The present invention also encompasses nucleotide sequences that hybridize under salt and temperature conditions equivalent to those described above to genomic DNA, plasmid DNA, cDNA, or synthetic DNA molecules that encode the same amino acid sequences as these nucleotide sequences, and genetically degenerate forms thereof due to the degenerancy of the genetic code, and that code on expression for a polypeptide that has the same or similar enzymatic activity as that of the polypeptides disclosed herein.

Biologically functional equivalent nucleotide sequences of the present invention also include nucleotide sequences that encode conservative amino acid changes within the amino acid sequences of the present polypeptides, producing silent changes therein. Such nucleotide sequences thus contain corresponding base substitutions based upon the genetic code compared to the nucleotide sequences encoding the present polypeptides. Substitutes for an amino acid within the fundamental polypeptide amino acid sequences discussed herein can be selected from other members of the class to which the naturally occurring amino acid belongs. Amino acids can be divided into the following four groups: (1) acidic amino acids; (2) basic amino acids; (3) neutral polar amino acids; and (4) neutral non-polar amino acids. Representative amino acids within these various groups include, but are not limited to: (1) acidic (negatively charged) amino acids such as aspartic acid and glutamic acid; (2) basic (positively charged) amino acids such as arginine, histidine, and lysine; (3) neutral polar amino acids such as glycine, serine, threonine, cyteine, cystine, tyrosine, asparagine, and glutamine; and (4) neutral nonpolar (hydrophobic) amino acids such as alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine.

Conservative amino acid changes within the present polypeptide sequences can be made by substituting one amino acid within one of these groups with another amino acid within the same group. The encoding nucleotide sequences (gene, plasmid DNA, cDNA, synthetic DNA, or mRNA) will thus have corresponding

base substitutions, permitting them to code on expression for the biologically functional equivalent forms of the present polypeptides.

Useful biologically functional equivalent forms of the DNA sequences disclosed herein include DNAs comprising nucleotide sequences that exhibit a level of sequence identity to corresponding regions or moieties of these DNA sequences from 40% sequence identity, or from 60% sequence identity, or from 80% sequence identity, to 100% sequence identity to the DNAs encoding the presently disclosed polypeptides. However, regardless of the percent sequence identity of these nucleotide sequences, the encoded proteins would possess the same or similar enzymatic activity as the present polypeptides. Thus, biologically functional equivalent nucleotide sequences encompassed by the present invention include sequences having less than 40% sequence identity to any of the nucleic acid sequences presented herein, so long as they encode polypeptides having the same or similar enzymatic activity as the polypeptides disclosed herein.

Sequence identity can be determined using the "BestFit" or "Gap" programs of the Sequence Analysis Software Package, Genetics Computer Group, Inc., University of Wisconsin Biotechnology Center, Madison, WI 53711.

Due to the degeneracy of the genetic code, i.e., the existence of more than one codon for most of the amino acids naturally occuring in proteins, genetically degenerate DNA (and RNA) sequences that contain the same essential genetic information as the DNA sequences disclosed herein, and which encode the same amino acid

sequences as these DNA sequences, are encompassed by the present invention. Genetically degenerate forms of any of the other nucleic acid sequences discussed herein are encompassed by the present invention as well.

The alternative nucleotide sequences described above are considered to possess a biological function substantially equivalent to that of the polypeptide-encoding DNAs of the present invention if they encode polypeptides having enzymatic activity differing from that of any of the present polypeptides by about 30% or less, preferably by about 20% or less, and more preferably by about 10% or less when assayed *in vivo* by complementation or *in vitro* by the standard enzymatic assays.

Example 2

Cloning and Sequencing of a cDNA Encoding the *Arabidopsis thaliana* Dihydrolipoamide S-acetyltransferase (E2) Component of the Plastid Pyruvate Dehydrogenase Complex

A search of the *Arabidopsis* expressed sequence tagged (EST) database identified one *Arabidopsis thaliana* EST clone which has significant homology to the (cyanobacterial) *Synechocystis* sp. dihydrolipoamide acetyltransferase subunit, GenBank accession D90915. The *Arabidopsis* EST clone (GenBank accession W43179) was obtained from the Arabidopsis Biological Resource Center (ABRC) at Ohio State University, then used to screen an *Arabidopsis* λPRL2 cDNA library (ABRC) for a full length clone as in Example 1. Two (approximately 1700 bp)

clones assessed as full length, were identified and sequenced as in Example 1.

The plastid PDC E2 clone is 1709 bp in length (SEQ ID NO:5; GenBank accession AF066079) with a continuous open reading frame of 1440 bp encoding a protein of 480 amino acids (SEQ ID NO:6), with a deduced molecular mass of 52,400 daltons. The mature portion of the E2 component, without the chloroplast targeting peptide (see below), has a deduced molecular mass of 44,900 daltons. When subjected to SDS-PAGE electrophoresis, the full length and the mature plastid PDC E2 proteins ran slower than a globular protein of the same mass. These proteins appeared on SDS-PAGE to have molecular masses of 69,000 and 62,000, respectively. This slow migration on SDS-PAGE electrophoresis is consistent with the electrophoretic behavior of mitochondrial E2 components (Guest et al., 1985).

The mature part of the cDNA clone (coding for the catalytic region of the protein) was expressed in *E. coli* using the pET28c expression vector (Novagen, Madison, WI). The recombinant protein (which includes a C-terminal six histidine tag) was purified under denaturing conditions by Ni-NTA affinity chromatography according to the manufacturer's instructions (Qiagen Inc., Chatsworth, CA). Polyclonal antibodies were raised to the recombinant protein in New Zealand White rabbits. These antibodies recognize the recombinant protein at a high dilution (1:100,000). In a analysis of an extract of purified pea chloroplasts, these antibodies recognized two proteins. One protein electrophoretically migrated at an apparent mass of 62,000, identical to the electrophoretic behavior of the mature plastid PDC E2 component. The other protein which was recognized by the

anti-E2 antibodies had an electrophoretic mobility with an apparent mass of 76,000 daltons. This larger protein is likely due to mitochondrial contamination, since its apparent mass is equivalent to the mitochondrial E2 component.

The cDNAs for the *Arabidopsis thaliana* plastid E1α, E1β, and E2 were transcribed and translated *in vitro* using the TnT™ transcription/translation system (Promega, Madison, WI) with the plasmid pZL1 (Life Technologies, Inc.) and the T7 promoter. Presenting the product to isolated pea chloroplasts resulted in ATP-dependent import into the plastid in a manner that protects it from protease action. This establishes that the cDNA sequences encode plastid targeting sequences. These targeting sequences are assessed to be the first 68 amino acids of the E1α subunit (Appendix B and SEQ ID NO:2), the first 73 amino acids of the E1ß subunit (Appendix D and SEQ ID NO:4), and the first 54 amino acids of the E2 component (SEQ ID NO:6).

Example 3

Cloning and Sequencing of cDNA Encoding the *Arabidopsis thaliana* E1α Subunit of the Branched-Chain Oxoacid Dehydrogenase Complex

Selection of an *A. thaliana* expressed sequence tagged (EST) cDNA clone (Newman et al., 1994) was accomplished by searching the Arabidopsis EST database using the BLASTP program of the National Center for Biotechnology Information. One EST cDNA clone (GenBank accession N96041) was found to have significant homology to the tomato, human, and bovine BCOADC E1α subunits, making it a candidate for the *A. thaliana* E1α. This cDNA

clone was obtained from the Arabidopsis Biological Resource Center at the Ohio State University. The clone was sequenced completely on both strands by subcloning restriction enzyme fragments of the clone and using two specific oligonucleotide primers designed from previously sequenced stretches. Sequencing was conducted by the DNA core facility at the University of Missouri, Columbia, MO on an ABI 377 instrument. The BCOADC E1α cDNA clone is 1587 bp, with a 3' untranslated region of 165 bp (Appendix E and SEQ ID NO:11). The open reading frame encodes a protein of 472 amino acids (Appendix F and SEQ ID NO:12) with a deduced molecular mass of 53,363 daltons. We have not identified an initiating methionine/start codon, but alignment with the tomato, bovine, human and mouse sequences shows the clone is considerably longer than the mature coding region of these proteins.

The deduced amino acid sequence of the clone has significant homology to BCOADC E1α sequences in the database: 56.8% identity with the tomato, 42% with the human, 40.7% with the bovine, and 41.6% with the mouse E1α amino acid sequences. Though an initiating methionine was not identified, the N-terminus has properties similar to a mitochondrial targeting peptide. The PSORT program (prediction of protein intracellular localization sites) suggests the mitochondrial matrix as the most probable destination of the *A. thaliana* E1α protein. However, the amino acid sequence also contains an SKL motif close to the C-terminus which is indicative of peroxisomal localization, and this is the second most probable localization site determined by the PSORT program.

Ser_{366} of the *A. thaliana* amino acid sequence is at a position which is conserved in all the above sequences. This site is a designated phosphorylation site for the mouse and bovine sequences. However, the second conserved Ser phosphorylation site in the animal sequences is replaced by a Pro in the tomato sequence and an Ala in the *A. thaliana* sequence (Appendix F and SEQ ID NO:12).

Example 4

Cloning and Sequencing of cDNA Encoding the *Arabidopsis thaliana* E1β Subunit of the Branched-Chain Oxoacid Dehydrogenase Complex

Selection of *Arabidopsis thaliana* expressed sequence tagged (EST) clones (Newman et al., 1994) was accomplished by searching the *Arabidopsis* EST database using the BLASTP PROGRAM of the National Center for Biotechnology Information. Two EST clones were found to have significant homology to the human and bovine branched-chain oxoacid dehydrogenase (BCOADC) E1β subunit. These two clones (GenBank accessions T04217 and H37020) were identified as potentially encoding the *Arabidopsis thaliana* BCOADC E1β subunits. We obtained these partial EST clones from the Arabidopsis Biological Resource Center (ABRC) at Ohio State University. One of these clones, GenBank accession T04217, was used to screen an *Arabidopsis* cDNA library for full length clones. The EST cDNAs were gel purified from low-melting agarose and probes prepared by labeling with [α^{32}P]dATP using a random prime oligonucleotide labeling kit (Pharmacia, Piscataway, NJ). Probes were desalted using Sephadex G-50 chromatography to remove unincorporated

nucleotides. An *Arabidopsis* cDNA library (λ-PRL2, obtained from the ABRC) was plated at a density of $2.9x10^4$ plaques per plate for a total of $2.03x10^5$ plaques. Biotrace NT nylon filters (Gelman, Ann Arbor, MI) were used for plaque-lifts and were processed according to the manufacturer's specifications. Prehybridization and hybridizations were performed according to *Current Protocols in Molecular Biology*, (Ausubel, et al., 1994). After three successive rounds of screening, 5 independent potential E1β cDNA clones were isolated, ranging in size from 500 to 1400 bp. Two of the five cDNA clones were selected for sequencing. Plaque-purified λ phage were treated according to the manufacturer's instructions (GibcoBRL, Gaithersburg, MD) in order to excise the pZL-1 recombinant clones. The cDNA sequences were obtained by sequencing both strands of the cDNA clone (and deletion fragments derived therefrom) using the Dye-deoxy terminating cycle sequencing reactions and an ABI prism Model 377 sequencer, according to the manufacuturer's instructions. Results from sequencing reactions were analyzed using IntelliGenetics GeneWorks DNA analysis program version 2.5 for Macintosh computers. Both cDNAs were identical. The BCOADC E1β cDNA is 1319 bp (Appendix G and SEQ ID NO:13) and contains a 133 bp 5′ untranslated region, an open reading frame of 1056 bp followed by 130 bp 3′ untranslated region. The open reading frame encodes a protein with 352 deduced amino acids (Appendix H and SEQ ID NO:14) with a calculated mass of 37,810 Daltons.

Fig. 10
~~Table 4~~ shows the alignment of the deduced amino acid sequences of various BCOADC E1β subunits. "." indicates conserved amino acids; "-" indicates a gap inserted to maximize homology. The deduced amino acid

sequence is 59% identical to the mammalian BCOADC E1β subunit (~~Table 4~~ Fig. 10). The primary sequence contains no obvious organellar targeting information.

The cDNA was expressed in *E. coli* after insertion into the plasmid vector pMal (New England Biolabs). The purified protein was used to prepare polyclonal antibodies which recognize the recombinant protein.

Example 5

Cloning and Sequencing of cDNA Encoding the *Arabidopsis thaliana* Dihydrolipoamide S-acyltransferase (E2) Component of the Branched-Chain 0xoacid Dehydrogenase Complex

A search of the *Arabidopsis* expressed sequence tagged (EST) database identified two *Arabidopsis thaliana* EST clones which have significant homology to the bovine and human branched-chain dihydrolipoamide acyltransferase subunit. These clones (GenBank accessions T42996 and N37840) were obtained from the Arabidopsis Biological Resource Center (ABRC) at Ohio State University. Sequencing of the 5' ends of the two clones showed only one to be a branched-chain E2 sequence (the other contained vector sequence only). The branched-chain EST clone (GenBank accession T42996) was sequenced completely on both strands by subcloning of restriction enzyme derived fragments and by primer walking. Sequencing reactions and analysis were performed as in Example 1.

The clone (SEQ ID NO:15) is 1618 bp in length and contains an open reading frame of 1449 bp encoding a protein of 483 amino acids (SEQ ID NO:16) with a predicted molecular mass of 52,729 daltons. Part of the cDNA clone (coding for the lipoyl and subunit-binding

domains, and part of the catalytic domain) was expressed in *E. coli* using the pET28a expression vector (Novagen, Madison, WI). The recombinant protein (which includes a C-terminal six histidine tag) was purified under denaturing conditions by Ni-NTA affinity chromatography according to the manufacturer's instructions (Qiagen Inc., Chatsworth, CA). Polyclonal antibodies were raised to the recombinant protein in New Zealand White rabbits. These antibodies recognize the recombinant protein at a high dilution (>1:100,000).

Example 6

Engineering Chimeric Branched Chain Oxoacid Dehydrogenase Complex E1α and E1β Subunits to Utilize the Plastid Pyruvate Dehydrogenase Complex E2 and E3 Components to Form a Hybrid Complex

The cDNA (or other encoding DNA) of the BCOADC E1β subunit can be used to form a chimeric protein targeted to the plastid to utilize the plastid pyruvate dehydrogenase complex (PDC) E2 component to produce propionyl-CoA. The chimeric BCOADC E1β subunit can be modified to comprise the E2 binding region of the plastid PDC E1β subunit and a plastid targeting sequence. The thus modified BCOADC E1β subunit can then be imported into the chloroplast, where it binds to the plastid PDC E2 component and, in conjunction with the plastid PDC E3 component, catalyzes the production of propionyl-CoA from 2-oxybutyrate. This leads to the production of the PHA precursor 3-hydroxyvaleryl-CoA, and consequently to biosynthesis of the PHA co-polymer poly(3HB-co-3HV) in plants that have been engineered to contain other enzymes

necessary for biosynthesis of this copolymer, as discussed above.

The nucleotide sequence that encodes the BCOADC E1β region 1 (the region or domain of the E1β protein that binds the BCOADC E1β component to the E2 core of the BCOADC complex [Wexler et al., 1991]) can be excised and replaced with the nucleotide sequence corresponding to the PDC E2 binding region from the plastid PDC E1β subunit (Johnston et al., 1997; Luethy et al., 1994). The construct can be further engineered to comprise a plastid targeting sequence of another plastid protein such as the Rubisco small subunit (Table 1) (von Heijne et al., 1991), or to comprise the plastid targeting sequence of the plastid PDC E1β subunit described by Johnston et al. (1997). See Figure 7B.

Chimeric fusions of plastid targeting sequences and the BCOADC E1α and E1β subunits can be generated by amplifying fragments of DNA coding for the regions involved. Chloroplast targeting peptides from each of the plastid PDC E1 subunits (PDC E1α and E1β) (Johnston et al., 1997) can be amplified from the original cDNAs (SEQ ID NOs 1 and 3). Similarly, the mature portions of the BCOADC E1α and E1β subunits can be amplified from their cDNAs (SEQ ID NOs 11 and 13). A unique restriction site can be included in the primer design to permit ligation of the chloroplast targeting peptides in-frame with the mature portions of the BCOADC E1α and E1β subunits.

To produce a BCOADC E1β chimera that can associate with the PDC E2 subunit, one can modify the BCOADC E1β subunit to include the plastid PDC E1β targeting peptide along with the plastid PDC E1β E2 binding region. In the final construct, the sequence for the E2 binding region

follows (i.e., is 3' to) the sequence for the targeting peptide, so that the chimeric BCOADC E1β protein contains approximately one-third plastid PDC E1β presequence (for example, amino acid residues 1 through 146 of SEQ ID NO:4) and the remainder consists of the BCOADC E1β subunit (for example, amino acid residues 94 through 352 of SEQ ID NO:14). The PDC E1β chloroplast targeting peptide and plastid PDC E2 binding region of the PDC E1β subunit can be amplified from the plastid PDC E1β cDNA (SEQ ID NO:4) using the following gene specific primer (SEQ ID NO:28) and a commercially available primer (e.g. M13/pUC forward primer, available from e.g. Stratagene, La Jolla, CA).

Forward oligonucleotide: 5′ GGGCCC CATATG TCTTCGATAATC 3′ (SEQ ID NO:28). Nucleotides 7 through 21 are preceded by an Nde1 enzyme site.

The mature part of the BCOADC E1β sequence (excluding the native BCOADC E2 binding site) can be amplified from the cDNA of SEQ ID NO:13 using the following gene specific primers:

Forward oligonucleotide: 5′ GGGCCC ACCGGT TTTGGCATTGGTCTA 3′ (SEQ ID NO:24). Nucleotides 406 through 423 are preceded by an Age1 enzyme site.

Reverse oligonucleotide: 5′ GGGCCC GAATTC TCATTACTAGTAATTCAC AGT 3′ (SEQ ID NO:25). Nucleotides 1177 through 1191 are preceded by an EcoR1 enzyme site.

The resulting truncated BCOADC E1β sequence can be ligated to the plastid PDC E1β sequence using the Age1 enzyme site already present in the plastid PDC sequence at a convenient position (amino acid residue 146). The above primers can be utilized to produce DNA fragments useful in joining the noted regions of the plastid PDC and BCOADC E1β sequences without any introduced or

substituted amino acids (Figure 7B).

To produce a BCOADC E1α chimera that can be targeted to a plastid, a chloroplast targeting peptide, for example the chloroplast targeting peptide from the plastid PDC E1α subunit (Johnston et al., 1997) (corresponding to amino acid residues 1 through 68) can be attached 5' to the mature portion of the BCOADC E1α subunit. A DNA fragment corresponding to the plastid targeting peptide can be amplified from the original PDC E1α cDNA (SEQ ID NO:1) using the following gene specific primers (SEQ ID NO:29 and SEQ ID NO:30):
Forward primer: 5' GGGCCC CCATGG CGACGGCTTTCGCT 3' (SEQ ID NO:29). Nucleotides 107 to 124 are preceded by an NcoI enzyme site.
Reverse primer: 5' GGGCCC TGATCA TATTATTGGTGGATTGCTT 3' (SEQ ID NO:30). Nucleotides 311 to 328 are preceded by a BclI enzyme site.

The entire mature coding region of the BCOADC E1α subunit can then be excised from the cDNA (SEQ ID NO:11) using convenient restriction enzyme sites, BclI at nucleotides 195 through 200, and XbaI at nucleotides 1424 through 1429. This includes the 3' stop codon.

The restriction enzyme fragments generated from both the plastid PDC and BCOADC E1α sequences can then be ligated together and subcloned into an appropriate vector (e.g. pZL1, Life Technologies Inc., Gaithersberg, MD). The BclI site used to ligate the two sequences introduces a single His residue between the plastid PDC E1β targeting peptide and the BCOADC E1α mature region.

The consequence of this addition can be determined experimentally to assess its impact, if any, on import and processing of the BCOADC E1α subunit, and on assembly of the hybrid BCOADC E1 complex.

An alternative approach to ligating the plastid PDC and BCOADC E1α sequences using the BclI site is to use a NotI site in its place in the design of the reverse oligonucleotide for the plastid targeting peptide, as follows (SEQ ID NO:19):

Plastid PDC E1α reverse primer: 5' GGGCCC GCGGCCGC ATTATTGGTGGATTGCTT 3' (SEQ ID NO:19). Nucleotides 311 through 328 are preceded by a NotI enzyme site.

The coding region for the mature BCOADC E1α protein (Appendix F and SEQ ID NO:12) can then be amplified from the cDNA (SEQ ID NO:11) using the following gene-specific primers:

Forward primer: 5' GGGCCC GCGGCCGC TGATCATTTGGTTCAGCAG 3' (SEQ ID NO:20). Nucleotides 195 through 213 are preceded by a NotI enzyme site.

Reverse primer: 5' GGGCCC GTCGAC TCAAACATGAAAGCCAGG 3' (SEQ ID NO:21). Nucleotides 1405 through 1422 are preceded by a SalI enzyme site and includes the stop codon.

Ligation of the two resulting sequences using the NotI enzyme site will introduce three Ala residues between them, which would overcome the introduction of a charged residue (His) using the BclI site described above.

To confirm the ability of the chimeric BCOADC E1α and E1β proteins to be imported into chloroplasts, the DNA encoding these chimeric proteins can be subcloned into a transcription vector such as pZL1 (Life Technologies Inc., Gaithersberg, MD) with the T7 promoter. The chimeric proteins are then transcribed/translated *in vitro*, for example using the TnT™ transcription/translation system (Life Technologies Inc.), and import assays with isolated chloroplasts can

be performed. This is a reliable assay to test the import and assembly of the chimeric proteins.

Experimental results have established that *in vitro* imported plastid PDC E1α and E1β subunit proteins associate to form the plastid pyruvate dehydrogenase heterotetramer within the chloroplast matrix, and that this heterotetramer associates with imported PDC E2 subunits (Randall et al., unpublished).

To obtain constitutive expression of the chimeric proteins in plants, their coding regions are preferably fused to the CaMV 35S promoter sequence. For dicotyledonous plants, the use of the pZP200 binary vector, for *Agrobacterium* transformation, is preferred.

The chimeric nucleic acids disclosed above are used to transform *Arabidopsis thaliana* or other plants by various methods well known in the art. As one alternative, the BCOADC E1α-chimeric construct comprising the plastid PDC E1α targeting sequence is used to produce transformed plants that are then crossed with plants that have been transformed with the BCOADC E1β-chimeric construct containing the plastid PDC E1β subunit targeting sequence and E2 component binding region.

As another alternative, a compound construct containing both the plastid-targeted BCOADC E1α-chimera and the plastid-targeted BCOADC E1β-chimera containing the PDC E1β E2 binding region is constructed in the form of a mega plasmid and used to transform plants by standard protocols for expression of both subunit chimeras simultaneously (Figure 7D). This can be achieved by including a stop signal at the 3' end of the BCOADC E1α chimeric sequence and a NOS transcription termination sequence. In order to obtain co-expression

of the two chimeric sequences, a second CaMV 35S promoter sequence can be placed 3' to the transcription termination sequence of the plastid-targeted BCOADC E1α chimeric coding sequence. This second promoter sequence can in turn be followed by the sequence coding for the BCOADC E1β chimera. This creates a mega plasmid or compound construct coding for both the BCOADC E1α and β subunit chimeras (Figure 7D).

The BCOADC E1α and β subunit chimeras thus targeted to the plastid bind to the plastid PDC E2 component (E2 components form the core of the complexes to which the E1 and E3 components bind). Since the chimeric BCOADC E1β subunit comprises the plastid PDC E1β E2 binding domain, a hybrid complex is formed. This hybrid complex is designed to have an enhanced ability to utilize 2-oxobutyrate as substrate in order to produce propionyl-CoA for 3-HV biosynthesis. Transgenic plants containing this hybrid complex can then be crossed by standard protocols with plants having enhanced ability to generate 2-oxobutyrate in the plastid compartment produced as described, for example, in Gruys et al. (1998).

Example 7

Targeting the BCOADC E1α, E1β, and E2 components to the Plastid to Form a Hybrid Complex with the Plastid PDC E3 Component

DNAs encoding the BCOADC E1α and β subunits and E2 component can be fused with plastid targeting sequences to direct importation of these proteins into the plastid to enhance propionyl-CoA production from 2-oxobutyrate. In this method, constructs of the BCOADC E1α and β subunits, the BCOADC E2 component, and, if desired, the

BCOADC E3 subunit, can be made with plastid targeting sequences, for example with plastid targeting sequences of the plastid pyruvate dehydrogenase complex (PDC) E1α and β subunits (Johnston et al., 1997) or the plastid PDC E2 component. See Figures 7A, 7C, and 7E. These constructs can be used to transform plants individually (followed by genetic crossing to combine the necessary components from each plant) or together to direct the desired BCOADC components to the plastid. The BCOADC E1α-chimera is as described above in Example 6. The BCOADC E1β-chimera containing the PDC E1β E2 binding region is also described in Example 6. When the plastid-targeted BCOADC E2 chimera is also employed (see below), the E2 binding region of the BCOADC E1β subunit need not be replaced with the plastid PDC E1β subunit E2 binding region. Instead, only the plastid PDC E1β targeting peptide is attached to the mature portion of the BCOADC E1β subunit (still retaining the native binding site for the BCOADC E2 component) (Figure 7E). This can be achieved by amplifying the appropriate regions of the PDC and BCOADC E1β cDNA sequences or other functionally equivalent DNA sequences. That portion of the cDNA coding for the plastid targeting peptide of the PDC E1β (amino acids 1 through 97) can be amplified from the cDNA (SEQ ID NO.:3) using the following gene specific primers. This amplified fragment includes a portion of the linker region between the targeting peptide and the E2-binding region.

Forward oligonucleotide: 5′ GGGCCC CATATG TCTTCGATAATC 3′ (SEQ ID NO:22). Nucleotides 7 through 21 are preceded by an Nde1 enzyme site.

Reverse oligonucleotide: 5′ GGGCCC CTCGAG ACCTTCCTGAAGAGC 3′ (SEQ ID NO:23). Nucleotides 277 through 297 are

preceded by an Xho1 enzyme site.

The mature portion of the BCOADC E1β sequence (including the native BCOADC E2 binding region), i.e., amino acid residues 45 through 349, can be amplified from the cDNA of SEQ ID NO:13 using the following gene specific primers:

Forward oligonucleotide: 5′ GGGCCC CTCGAG ATCGCTTTGGACACC 3′ (SEQ ID NO:31). Nucleotides 262 through 277 are preceded by an Xho1 enzyme site.

Reverse oligonucleotide: 5′ GGGCCC GAATTC TCATTACTAGTAATTCAC AGT 3′ (SEQ ID NO:25). Nucleotides 1177 through 1191 are preceded by an EcoR1 enzyme site.

Use of the foregoing oligonucleotide primers allows the joining of the appropriate plastid PDC and BCOADC E1β sequences without any introduced or substituted amino acids (Figure 7E). As disclosed in Example 6, the resulting DNA can be subcloned into a transcription vector to test import and assembly prior to transformation of *Arabidopsis* or other plants (or prior to the construction of a mega plasmid for co-expression, cf. Figure 7D).

Further to the above, a chimera comprising the plastid targeting sequence (nucleotides 59-232) of the plastid PDC E2 (dihydrolipoamide acetyltransferase) component and the sequence for the mature BCOADC dihydrolipoamide acyltransferase (E2) subunit can be constructed. The N-terminus of the BCOADC E2 subunit can be replaced with the chloroplast targeting peptide from the plastid PDC E2 subunit. In this case, the native E2 binding domain of the BCOADC E1β subunit need not be replaced with the E2 binding domain of the plastid PDC E1β subunit as described in Example 6. Only the plastid PDC E2

targeting peptide is needed because the BCOADC E2 component which is imported into the plastid will naturally associate with the BCOADC E1β subunit.

The plastid targeting sequence can be amplified from the plastid PDC E2 cDNA of SEQ ID NO:5 using the following gene-specific primers:
Forward primer: 5′ GGGCCC CATATG GCGGTTTCTTCT 3′ (SEQ ID NO:26). Nucleotides 59 through 73 are preceded by an Nde1 enzyme site.
Reverse primer; 5′ GGGCCC CCATGGC AATTTCAGGATTCTT 3′ (SEQ ID NO:27). Nucleotides 218 through 232 are preceded by an Nco1 enzyme site.

The region coding for the mature portion of the BCOADC E2 protein can be excised from the cDNA (SEQ ID NO.:15) using convenient restriction enzymes (Nco1 and Not1). This DNA fragment is then ligated in-frame with the PDC E2 plastid targeting peptide using the common Nco1 enzyme site (Figure 7C). As described in Example 6, the import and assembly of this chimeric E2 subunit can be examined by *in vitro* import assays. Efficient import of the BCOADC E2 protein into isolated pea chloroplasts and formation of a complex with both the endogenous PDC heterotetramer and imported BCOADC E1α-E1β heterotetramer can be determined.

The plastid-targeted branched-chain oxoacid dehydrogenase complex components utilize any 2-oxobutyrate (α-ketobutyrate) produced in the plastid to make propionyl CoA, which in turn is a substrate for the enzymes producing polyhydroxyalkanoic acids (PHAs).

As previously indicated, it appears to be unnecessary to prepare a plastid-targeted construct

for the BCOADC E3 component since the E3 components of all of the mitochondrial α-ketoacid dehydrogenase complexes appear to be interchangeable. The PDC E3 component already present in the plastid should function with the plastid-targeted BCOADC E1α, E1β, and E2 subunits. If desired, one can, for example, place a plastid targeting sequence on the mitochondrial E3 component in place of the first 31 amino acids of the mitochondrial PDC E3 reported by Turner et al. (1992) (GenBank accession number X2995), corresponding to the first 72 nucleotides of that particular cDNA. This is done by standard protocols well known to those skilled in the art.

As discussed above, the plastid is capable of PHA biosynthesis when the appropriate enzymes are present in the plant (Poirier et al., 1992; Nawrath et al., 1994). Targeting BCOADC subunits and components to this organelle as described in Examples 6 and 7 herein further enhances ability of plants to biosynthesize the 3HB-co-3HV copolymer.

The invention being thus described, it will be obvious that the same can be varied in many ways. Such variations are not to be regarded as a departure from the spirit and scope of the present invention, and all such modifications and equivalents as would be obvious to one skilled in the art are intended to be included within the scope of the following claims.

References Cited

Ainley et al., (1990) Plant Mol. Biol. 14: 949.

Ali et al., (1993) J. Biol. Chem. 268: 22353.

Ausubel et al., (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc.

Ausubel et al., (1994) *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc.

Back et al., (1991) Plant Mol. Biol. 17: 9.

Becker et al., (1994) Plant J. 5: 299.

Behal et al., (1989) Biochem. Biophys. Res. Commun. 164: 941.

Bower and Birch, (1992) Plant J. 2: 409.

Brandl et al., (1990) Adv. Biochem. Eng. Biotech. 41: 77.

Budde et al., (1991) Plant Physiol. 95: 131.

Bustos et al., (1991) EMBO J. 10: 1469.

Byrom, (1987) Trends Biotechnol. 5: 246.

Bytebier et al., (1987) Proc. Natl. Acad. Sci. USA 84: 5345.

Camp and Randall, (1985) Plant Physiol. 77: 571.

Camp et al., (1988) Biochim. Biophys. Acta 933: 269.

Cassas et al., (1993) Proc. Natl. Acad. Sci. USA 90: 11212.

Castresana et al., (1988) EMBO J. 7: 1929.

Christou (March/April, 1994) Agro Food Industry Hi Tech, p.17.

Christou et al., (1991) Bio/Technology 9: 957.

Dawes and Senior, (1973) Adv. Microb. Physiol. 14: 135.

De la Pena et al., (1987) Nature 325: 274.

Doi, (1990) *Bacterial Polyesters*, VCH Publishers, Inc., New York.

Doyle et al., (1986) J. Biol. Chem. 261: 9228.

Eswaran et al., (1995) Biochim. Biophys. Acta 1252: 203.

Feinbaum et al., (1991) Mol. Gen. Genet. 226: 449.

Fisk and Dandekar, (1993) Scientia Horticulturae 55: 5.

Fosket, (1994) *Plant Growth and Development*, Academic Press, Inc., San Diego, p. 132.

Fromm et al., (1990) Bio/Technology 8: 833.

Gasser and Fraley, (1989) Science 244: 1293.

Gavel and von Heijne, (1990) FEBS 261: 455.

Gerbling and Gerhardt, (1988) Plant Physiol. 88: 13.

Gerbling and Gerhardt, (1989) Plant Physiol. 91: 1387.

Gopalakrishnan et al., (1989) Biochem. Biophys. Res. Commun. 160: 715.

Gordon-Kamm et al., (1990) Plant Cell 2: 603.

Grof et al., (1995) Plant Physiol. 108: 1623.

Gruys et al., (1998) PCT International Publication WO 98/00557.

Guan et al., (1995) J. Biol. Chem. 270: 5412.

Guest et al., (1985) J. Molec. Biol. 185:743.

Hoffman, U.S. Patent No. 5,106,739

Horn et al., (1988) Plant Cell Rep. 7: 469.

Johnson et al., (1992) Molec. Biochem. Parasitol. 51: 37.

Johnston et al., (1997) Biochim. Biophys. Acta. 1321: 200.

Jones and Yeaman, (1986) Biochem. J. 237: 621.

Kang and Rawsthorne, (1994) Plant J. 6: 795.

Kaneko et al., (1996) DNA Research 3: 109.

Karchi et al., (1993) Plant J. 3: 721.

Kares et al., (1990) Plant Mol. Biol. 15: 905.

Knutzon et al., (1992) Proc. Natl. Acad. Sci USA 89: 2624.

Korotchkina and Patel, (1995) J. Biol. Chem. 270: 14397.

Koziel et al., (1993) Bio/Technology 11: 194.

Kridl et al., (1991) Seed Sci. Res. 1: 209.

Kuhlemeier et al., (1989) Plant Cell 1: 471.

Layman, (Oct. 31, 1994) Chem. & Eng. News, p. 10.

Lam and Chua, (1990) Science 248: 471.

Lemoigne, (1926) Bull. Soc. Chim. Biol. (Paris) 8: 770.

Lindsay, (1992) Modern Plastics 2: 62.

Luethy et al., (1994) Biochim. Biophys. Acta 1187: 95.

Luethy et al., (1995) Gene 164: 251.

Luethy et al., (1995) J. Plant Physiol. 145: 443.

Maliga et al., U.S. Patent No. 5,451,513.

Maliga et al., PCT International Publication WO 95/16783.

Maliga et al., PCT International Publication WO 95/24492.

Maliga et al., PCT International Publication WO 95/24493.

Miernyk and Dennis, (1983) J. Exp. Bot. 34: 712.

Minsky et al., (1997) Abst. 5th Int. Cong. Plant Molec. Biol., Abstract No. 302.

Nakamura et al., (1992) Int. J. Biol. Macromol. 14: 321.

Nawrath et al., (1994) Proc. Natl. Acad. Sci *USA* 91: 12760.

Newman et al., (1994) Plant Physiol. 106: 1241.

Odell et al., (1985) Nature 313: 810.

Ohlrogge et al., (1979) Proc. Natl. Acad. Sci. USA 76: 1194.

Ou-Lee et al., (1986) Proc. Natl. Acad. Sci USA 83: 6815.

Padgette et al., (1995) Crop Sci. 35: 1451.

Palmer, (1992) in *Plant Gene Research: Cell Organelles* (Herrmann, R.G., ed.), Vol. 7, pp. 99-133, Springer-Verlag/Wien, Wurzburg.

Patel, M.S. and Roche, T.E. (1990) FASEB J. 4: 3224.

Paxton et al., (1986) Biochem. J. 234: 295.

Peoples and Sinskey, (1989a) J. Biol. Chem. 264: 15293.

Peoples and Sinskey, (1989b) J. Biol. Chem. 264: 15298.

Pettit et al., (1978) Proc. Natl. Acad. Sci. USA 75: 1881.

Poirier et al., (1992) Science 256: 520.

Poirier et al., (1995) Bio/Technology 13: 142.

Potrykus, (1991) Annu. Rev. Plant Physiol. Plant Mol. Biol. 42: 205.

Randall et al., (1977) Arch. Biochem. Biophys. 178: 342.

Randall et al., (1989) Ann. N.Y. Acad. Sci. 573: 192.

Randall et al., (1996) Proc. Phytochem. Soc. Europe 39: 87.

Reed, L.J. (1974) Acc. Chem. Res. 7: 40.

Reid et al., (1975) Biochem. Biophys. Res. Commun. 62: 42.

Reid et al., (1977) Plant Physiol. 59: 842.

Reith and Munholland, (1995) Plant Molec. Biol. Rep. 13: 333.

Rhodes et al., (1988) Science 240: 204.

Richins et al., (1987) NAR 20: 8451.

Sambrook et al., (1989) *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Schubert et al., (1988) J. Bacteriol. 170: 5837.

Schulze-Lefert et al., (1989) EMBO J. 8: 651.

Slater et al., (1988). J. Bacteriol. 170: 4431.

Slighton and Beachy, (1987) Planta 172: 356.

Smith et al., (1992) Plant Physiol. 98: 1233.

Somers et al., (1992) Bio/Technology 10: 1589.

Staub and Maliga, (1993) EMBO J. 12: 601.

Stayton et al., (1991) Aust. J. Plant. Physiol. 18: 507.

Steinbüchel, (1991) In: *Biomaterials*, D. Byrom (Ed.), MacMillan, Basingstoke, pp. 123-213.

Steinbüchel and Pieper, (1992) Appl. Microbiol. Biotechnol. 37: 1.

Svab et al., (1990) Proc. Natl. Acad. Sci. USA 87: 8526.

Svab and Maliga, (1993) Proc. Natl. Acad. Sci. USA 90: 913.

Taylor et al., (1992) Planta 188: 255.

Thompson et al., (1977a) Plant Physiol. 59: 849.

Thompson et al., (1977b) Plant Physiol. 59: 854.

Toriyama et al., (1988) Bio/Technology 6: 10.

Turner et al., (1992) J. Biol. Chem. 267: 7745.

Vasil et al., (1992) Bio/Technology 10: 667.
Von Heijne et al., (1989) Eur. J. Biochem. 180: 535.

Von Heijne et al., (1991) Plant Mol. Biol. Rep. 9: 104.

Wan and Lemaux, (1994) Plant Physiol. 104: 37.

Wang et al., (1992) Bio/Technology 10: 691.

Weeks et al., (1993) Plant Physiol. 102: 1077.

Weisshaar et al., (1991) EMBO J. 10: 1777.

Wexler et al., (1991) FEBS Lett. 282: 209.

Williams et al., (1979) Plant Physiol. 64: 1099.

Wynn et al., (1996) In: *Alpha-Keto Acid Dehydrogenase Complexes*. M.S. Patel, T.E. Roche and R.A. Harris (eds). Birkhäuser Verlag, Basel, Switzerland. pp. 101-117.

Yamaguchi-Shinozaki et al., (1990) Plant Mol. Biol. 15: 905.

Yeaman, (1989) Biochem. J. 257: 625.

Yoon et al., (1995) J. Ferment. Bioeng. 80: 350.

Zhang and Wu, (1988) Theor. Appl. Genet. 76: 835.

Zhong et al., (1993) Plant Cell Rep. 13: 1.

Zilkey and Canvin, (1969) Biochem. Biophys. Res. Commun. 34: 646.